ABSTRACT

TZENG, SZ-TING. Emotion and Norm Awareness Cognitive Systems. (Under the direction of Co-chair 1 of Committee and Co-chair 2 of Committee.)

With advances in technology, Artificial Intelligence (AI) has become part of our daily lives. Different from the past, the software is no longer limited to confined and isolated environments. Nowadays, software interacts with its environment but also with each other and with humans. With humans-in-the-loop, there are emerging needs for human factors to be considered when building modern AI systems. Specifically, these AI systems should reason over humans' behaviors determined by internal attitudes and external factors. Human values help to explain behaviors and attitudes from a motivational basis. Specifically, human values define an individual's intrinsic motivation and dominate how this individual thinks and evaluates everything. AI that considers human values and sociotechnical aspects in decision-making would be more realistic and trustworthy.

Another concern that arises from socio-technical systems is adaptability. In human societies, norms are expectations of individual behaviors, which can be established or revised from the top-down or emerge from the bottom-up. Norms from the top-down approach, such as laws, are defined by a centralized authority. On the contrary, norms can also emerge from the bottom up via agent interactions. In both approaches, norms and the environment can change over time. While introduced to multiagent systems (Normative MAS), social norms act as behavioral constraints that regulate agent behaviors within MAS. To reduce human interventions, the ability of adaptation for AI systems becomes necessary.

By regulating agent interactions, norms facilitate coordination in MAS. Sanctions, the reactions to norm satisfaction or norm violation, have guided research on norms for a long time. Current research on norms focuses on how sanctioning shapes agents' behaviors, specifically on the santionees. Little research on norm emergence has incorporated values, which determine how agents behave and how they assess each other's behavior. Sanctions in the real world are often more subtle than mere rewards or punishments. In particular, verbal messages or expressed emotions also serve as sanctions. We include emotions into the normative reasoning process, which evaluate whether to comply or violate norms. In addition, we consider expressed emotions as information. Normative information can be conveyed in explicit messages or emotional expressions, which other works have not considered. While humans evaluate social norms based on their values, they are flexible in accepting exceptions. Previous work on conflict resolution reveal all information to others. We propose work on providing deliberate justifications.

Emotion and Norm Awareness Cognitive Systems

by
Sz-Ting Tzeng

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

program name

Raleigh, North Carolina

2021

APPROVED BY:

Arnav Jhala	Min Chi
William Rand	Nirav Ajmeri

Munindar P. Singh
Chair of Advisory Committee

ACKNOWLEDGEMENTS

Thanks to family, advisor, committee, labmates, friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

With advances in technology, Artificial Intelligence (AI) has become part of our daily lives. e.g., Netflix's recommendation system suggests videos based on users' preferences; virtual assistants on smart devices process and execute your requests in natural language. Different from the past, software is no longer limited to confined and isolated environments. With modern technology, software interacts with its environment but also with each other and with humans (Kafalı et al. 2016). With humans-in-the-loop, there are emerging needs for human factors to be considered when building modern AI systems. Specifically, these AI systems should be able to reason over humans' behaviors determined by internal attitudes and external factors. AI that considers human values and sociotechnical aspects in decision-making would be more realistic and trustworthy.

In Attribution theory, internal attributions explain human behavior with a focus on the characteristics of a person (Gerace 2020). e.g., their personalities, abilities, and physical characteristics. On the contrary, external attributions stress environmental or situational factors. e.g., social influences and task difficulty. In the theory of basic human values, values characterize individuals and societies (Schwartz 2012). Values help to explain behaviors and attitudes from a motivational basis. Specifically, human values define the intrinsic motivation of an individual and dominate how this individual thinks and evaluates everything. When the socio-technical systems become more interconnected, the complexity of interactions increases drastically and become hard to model all the possibilities. Basing on values provides a solution to simplify the complexity.

Another concern that arises from the increasing size of the socio-technical systems is adaptability. In human societies, norms are expectations of individual behaviors (Savarimuthu and Cranefield 2011), which can be established or revised from the top-down or emerge from the bottom-up (Morris-Martin et al. 2019). Norms from the top-down approach, such as law or regulations, are expensive and defined by a centralized authority. On the contrary, norms can also emerge from the bottom up via agent interactions. In both approaches, norms and the environment can change over time. While introduced to multiagent systems (Normative MAS), social norms act as behavioral constraints that regulate agent behaviors within multiagent systems (Hollander and Wu 2011). To reduce human interventions, the ability of adaptation for AI systems becomes necessary.

While humans' decision-making includes internal and external attitudes, the other key factor in decision-making is emotions. Emotions, the responses to internal or external events or objects, may influence the decision-making and provide extra information in communication. Even more, emotions could be part of norms themselves. Including both norms and emotions helps us to build explainable and trustworthy AI.

In our attempt to include adaptation and explainability within AI systems, we consider normative MAS along with human values. We first focus on elements that help to regulate human behaviors. In this dissertation, we study how to endow a normative MAS with the ability to adapt to a dynamic environment and reason over human values.

1.1 Motivations

1.1.1 Emotion as sanctions

In multiagent systems, norms and sanctions are often used to regulate agent behaviors while maintaining their autonomy. However, sanctions in the real world are more subtle instead of harsh punishment. For instance, the sanctions could be trust update or emotional expression and might change one's behavior (Nardin et al. 2016; Bourgais et al. 2019). At the basic level of Emotions' Social Functions, emotions help individuals understand others' preferences, beliefs, and intentions and coordinate social interactions (Keltner and Haidt 1999).

Consider a pandemic scenario. During a pandemic, many stores limit the number of customers in stores at once to protect their customers. A side effect of this practice is the long queue outside the stores. While there is a social norm that people should line up to enter the stores, some can still jump the queue to get services in advance. Suppose those who violate the norms would feel guilty (self-directed emotion) and receive negative emotions from others (other-directed emotion). These felt emotions will enforce the norm in stores.

The above scenario demonstrates the necessity of incorporating emotions when studying norm enforcement.

1.1.2 Expressed emotion as information

Apart from being a way to sanction or being a factor to explain some human behaviors, expressed emotions as information enable inference of mental states that are otherwise not observable (Wu et al. 2018a; Wu and Schulz 2020). Compared to explicit expression, emotions show subtle normative information over behaviors. With explicit messages, humans gain direct and indirect normative information. An example of indirect messages is folklore in many countries. Kids should behave themselves to get rewards from Santa Claus. While messages provide clear normative information, emotions give subtle normative information for our behavior as well. Upon receiving negative emotions after some actions, we can infer that our behaviors do not fit into others' expectations.

Consider the following example. David notices Bella's suspicious symptoms and expresses little negative emotions with facial expressions, e.g., anger. Upon perceiving the emotion, Bella might infer the emotion was for her violation of self-quarantine and believes that some potential punishments may happen. Bella then changes her behavior. Other people who observe this may make the same inference and learn the causality.

The above example shows the other functionality of emotions than sanctions. Therefore, we found the necessity of considering emotions as information when studying norm adaptation. enforcement.

While previous works on normative MAS assumed regimented environments, earlier works on emergent behavior with reinforcement learning did not consider the concept of norms in the agent structure. To take advantage of the nature of reinforcement learning, trial and learn, and reduce human intervention, we adopt reinforcement learning. We show that reinforcement learning has the potential to model norms and emotions.

1.1.3 Values and social value orientation as preferences

While social norms regulate human behaviors, humans evaluate social norms based on human values and decide whether to comply or not. Consider there is a rare case scenario. Felix, who has always been a good citizen, finds someone breaks into his house and causes a safety threat. Being aware of the crime of assault, Felix chooses to overpower the suspect. Values may vary from person to person. A trustworthy AI system has to consider its stakeholder's values to make the right decision.

Another orientation an AI system needs to consider is social value orientation (SVO). Elliot, an advocate of individualism who cares only about his pleasure, goes to the nightclub with Debra and Cecilia while they know some confirmed cases showed up before. Although Elliot knows Debra values safety most, he chooses not to wear a mask to maximize his pleasure. As a prosocial person, Cecilia evaluates the values of everyone and chooses to wear the mask to achieve greater safety and satisfaction between herself and others.

Above scenarios demonstrate the necessity of incorporating values when building trustworthy and explainable AI.

1.2 Research Objectives

The research objective that we aim to achieve is to design a framework that ensures trustworthy AI systems that consider human values and operate in dynamic environments.

Based on aforementioned challenges and objectives, we seek to address the following questions.

$\text{RQ}_{\text{emotion}}$. How does modeling the emotional responses of agents to the outcomes of interactions affect the norm emergence?

To address $\text{RQ}_{\text{emotion}}$ that corresponded to Section 1.1.1, we refine the abstract normative emotional agent architecture (Argente et al. 2020) and investigate the interplay of norms and emotions. We propose a framework *Noe* based on BDI architecture (Bratman 1987), norm life-cycle (Savarimuthu and Cranefield 2011; Frantz and Pigozzi 2018; Argente et al. 2020), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009).

With *Noe*, when agent A violate a norm, agent A feels guilty about its behavior. In the meantime, agent B sanctions agent A for A's bad action. Note that sanctioning here is exclusively via emotions During interactions, agent A learns the norm from this experience.

To address the problems in Section 1.1.2, we investigate the following research questions.

$\mathrm{RQ}_{\mathrm{RL}}$. How does reinforcement learning accommodate reasoning about cognitive constructs, emotions, and norms?

$\mathrm{RQ}_{\mathrm{information}}$. How does providing indirect information, e.g., emotion as information, influence norm emergence?

We present an agent architecture that integrates reinforcement learning with normative reasoning, cognitive architecture, and emotion modeling to answer our first research question $\mathrm{RQ}_{\mathrm{RL}}$.

To address $\mathrm{RQ}_{\mathrm{information}}$, we define two expressions: explicit normative explanation (Andrighetto et al. 2013) and emotion as information. We apply belief reward shaping (Marom and Rosman 2018), a reward augmentation framework that considers rewards from the environment and also from beliefs, in our simulation. To answer $\mathrm{RQ}_{\mathrm{information}}$, we design a simulation with reinforcement learning (RL) and *Hermione*. Our proposed framework *Hermione* combines a cognitive architecture (simplified from the Belief-Desire-Intention (BDI) architecture (Rao and Georgeff 1991)), norm life-cycle (Argente et al. 2020; Frantz and Pigozzi 2018; Savarimuthu and Cranefield 2011), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009).

With *Hermione*, agent B can punish agent A for A's bad action where the punishment causes some loss to agent A. Agent B can instead send messages or express negative emotions to agent A. Agent A may associate the message or bad emotions with a potential penalty. With shared parameters, agents are homogeneous and assumed to share information and knowledge with each other (Liu 2020). Agent C, in the same society, observes B-A interacting and learns the norm.

To tackle the challenges in Section 1.1.3, we investigate the following research questions.

$\mathrm{RQ}_{\mathrm{Values}}$. How does basic human values influence the emergence of norms?

$\mathrm{RQ}_{\mathrm{SVO}}$. How does social value orientation influence the emergence of norms?

To address $\mathrm{RQ}_{\mathrm{Values}}$, we consider basic human values within agents. These values provide a basis to evaluate the world. To address $\mathrm{RQ}_{\mathrm{SVO}}$, we apply social preferences, social value orientation, along with basic human values. Our proposed framework *Fleur* combines world model, cognitive architecture, and social model. Both human values and social value orientation provide subjective weights on state evaluation.

With *Fleur*, agents consider their basic values and SVO to make decisions. In this approach, agents sanctions based on majority behaviors.

1.3 Contributions

1.3.1 Status and Tentative Timeline

Table 1.1: Proposed plan.

Project	**Status**	**Venue**	**Timeline**
$\text{RQ}_{\text{emotion}}$. *Noe* (Sanctioner's emotions; guilt)	Accepted	COINE at AAMAS 2021 (short)	–
RQ_{RL} and $\text{RQ}_{\text{information}}$. *Hermione* (Emotion for sanctioning and for information; third-party observation)	Pending	JAAMAS (in-review)	–
$\text{RQ}_{\text{Values}}$ and RQ_{SVO}. Values and social value orientation as preferences	In progress	JAIR or AIJ or JAAMAS	Sep 21 – Nov 21
RQ5. Deliberate justification	Early stage	IJCAI 2022 (due Jan 2022)	Nov 21 – Jan 22

1.4 Organization

The dissertation is organized as follows. Chapter two reviews the related works and backgrounds used in this dissertation. Chapter three presents our *Noe* framework to address $\text{RQ}_{\text{emotion}}$. Chapter four proposes our *Hermione* framework to tackle RQ_{RL} and $\text{RQ}_{\text{information}}$. Chapter four presents our *Fleur* framework to address $\text{RQ}_{\text{Values}}$ and RQ_{SVO}.

CHAPTER

2

LITERATURE REVIEW

2.1 Agent Architectures

Ortony et al. (1988) model emotions based on events, action, and objects. Marsella and Gratch (2009) propose a computational model of emotion to model appraisal in perceptual, cognitive, and behavioral processes. Wu et al. (2018b) adapt the classic Sally-Anne task (Wimmer and Perner 1983; Baron-Cohen et al. 1985), an experiment that demonstrates at what age children be able to form false beliefs. Wu's results show that young children can use others' emotional expressions to infer their behavior and knowledge. Moerland et al. (2018) survey computational models of emotions in reinforcement learning while focusing on agent emotions.

Argente et al. (2020) propose an abstract normative emotional agent architecture, an extension of BDI architecture that combines emotional, normative, and cognitive component. Argente et al. (2020) define four types of relationships between emotions and norms: (1) emotion in the process of normative reasoning, (2) emotion generation with norm satisfaction or violation, (3) emotions as a way to enforce norms, (4) anticipation of emotions promotes internalization and compliance of social norms.

Bourgais et al. (2019) present an agent architecture that integrates cognition, emotions, emotion contagion, personality, norms, and social relations to simulate humans and ensure explainable behaviors. In the proposed architecture, agents work in four steps: (i) agent A perceives the environment and links the environment and its knowledge; (ii) agent A updates its cognitive mental state, generates emotions based on its knowledge, and updates the social relation with others; (iii) agent A makes decisions via its cognitive engine and its normative engine; (iv) agent A generates emotions via its emotion model and applies the generated emotions to later decision making. (iv) the architecture generates a temporal dynamic by degrading the cognitive mental states and emotion's intensity and updates norm status. Emotions not only enforce social norms but also serve as activation of norms in this work.

von Scheve et al. (2006) allow emotion generation with norm satisfaction or violation. Specifically, an observer agent perceives the transgression of a norm of another, its strong negative emotions (e.g., contempt, disdain, detestation, or disgust) constitute negative sanctioning of the violator. The negative sanctioning then leads to negative emotions (e.g., shame, guilt, or embarrassment) in the violator. Besides, compliance with the social norms can stem from the fear of emotional-driven sanctions, which would lead to negative emotions in the violator. Such fear enforces social norms.

Tzeng et al. (2021) combine normative model, a BDI model, and emotions for the decision-making process.

2.2 Cognitive Architecture

In terms of intelligent agents, Broersen et al. (2001) introduce the Beliefs-Obligations-Intentions-Desires (BOID) architecture, where the BOID further includes obligation and conflict resolution. Andrighetto et al. (2013) show that a combination of verbal normative information, specifically positive normative content, and material punishment leads to higher and more stable cooperation with human subjects and agent-based simulation. These models include normative reasoning but leave out emotions. Kalia et al. (2019) demonstrate how emotions influence norm outcomes. Specifically, they consider norm outcomes with respect to emotions and trust and goals.

2.3 Norms and Agent-Based Simulation

Computer simulation has become an important tool in scientific research to model a variety of phenomena, e.g. physics, urban growth, aerospace, healthcare, etc. Among varied simulation methods, agent-based modeling and simulation provides a different approach, specifically emergent interactions, to model phenomena (Law et al. 2000). Hollander and Wu (2011) introduce social norms and its applications to agent-based system.

Dignum et al. (2020) associate the interventions that governments can take and their economic and social consequences with the SEIR model since effective and sustainable solutions cannot exist without considering these factors.

de Mooij et al. (2021) develop a large-scale data-driven agent-based simulation model where each agent reasons about their internal attitudes and external factors to simulate behavioral interventions in the real world.

Dell'Anna et al. (2020) introduce a norm revision component that uses data collected from interactions and an estimation of agents' preferences to modify sanctions at runtime.

Al-Amin et al. (2018) study individual and group sanctions to ensure compliance. In the meantime, other individuals who witness such violations and sanctions learn the relationship between that violation and negative sanctions.

2.4 Norm Emergence and Norm Robustness

Morris-Martin et al. (2019) provide a survey on norm emergence. They describe norms as expected behaviour in a society or behavior that is common in a society. Norm emergence is a bottom-up approach where agents learn the common behavior from interactions.

Mukherjee et al. (2008) investigate the effects of heterogeneous agents using different learning algorithms. In addition, they study norm emergence when agent interactions are physically constrained. Sen and Airiau (2007) propose a model that supports the emergence of social norms by learning from interactions.

Hao et al. (2017) propose two learning strategies based on local exploration and global exploration to support the emergence of social norms. The local and global learning strategies enable them to maximize the average payoffs among agents.

CHAPTER

3

A FRAMEWORK FOR EMOTIONAL NORMATIVE AGENT

3.1 Introduction

Humans, in daily life, face many choices at many moments, and each selection brings positive and negative payoffs. In psychology, decision-making (Simon 1960) is a cognitive process that selects a belief or a series of actions based on values, preferences, and beliefs to achieve specific goals. The final choice could be rational or irrational in terms of utility theory since emotions often influence decision-making (Schwarz 2000). Emotions, the responses to internal or external events or objects, can involve the decision-making process and provide extra information in communication (Keltner and Haidt 1999; Schwarz 2000). Social norms regulate behaviors in a human society (Singh 2013; Savarimuthu and Cranefield 2011), but humans and agents have to deviate from norms in certain contexts. For instance, people do handshaking at the normal time. Yet, people have to deviate from the social norm during a pandemic.

An agent that models the emotions of its users and other humans can potentially behave in a more realistic and trustworthy manner. The decision-making process for humans or agents involves evaluating possible consequences of available actions and choosing the action that maximizes the expected utility (Edwards 1954). Without considering emotions or other affective characteristics, such as personality or mood, some compliance seems irrational (Argente et al. 2020). Humans' compliance does show hints on rational planning over their objectives (Keltner and Haidt 1999). Including emotion or personality in normative reasoning makes these compliance behaviors explainable. Norms either are defined in a top-down manner or emerge in a bottom-up manner (Savarimuthu and Cranefield 2011). Works on norms include norm emergence based on the prior outcome of norms, automated run-time revision of sanctions (Dell'Anna et al. 2019), or considering various aspects during reasoning (Ajmeri et al. 2020, 2018). However, sanctions in the real world are more subtle instead of harsh punishment. For instance, the sanctions could be trust update or emotional expression and might change one's behavior (Nardin et al. 2016; Bourgais et al. 2019). Kalia et al. (Kalia et al. 2019) considered norm outcome with respect to emotions and trust and goals. Modeling and reasoning about emotions and other affective characteristics in an agent then become important in decision making and would help the agent enforce and internalize norms, which boost agent autonomy.

At the basic level, emotions help individuals understand others' preferences, beliefs, and intentions and therefore coordinate social interactions. Furthermore, emotions serve as motivations or deterrents for others' social behavior and therefore play an essential role in learning. At the cultural level, emotions facilitate cultural identities and help individuals learn the norms and the values of their culture (Keltner and Haidt 1999). Based on this understanding of emotions and norms, we propose an agent architecture that integrates the BDI architecture (Bratman 1987) with a normative model (Argente et al. 2020; Singh 2013) and emotional model (Alfonso Espinosa 2017; Marsella and Gratch 2009).

Accordingly, we investigate the following research question.

$\text{RQ}_{\text{emotion}}$. How does modeling the emotional responses of agents to the outcomes of interactions affect the norm emergence and social welfare in an agent society?

To address $\text{RQ}_{\text{emotion}}$, we refine the abstract normative emotional agent architecture (Argente et al. 2020) and investigate the interplay of norms and emotions. We propose a framework *Noe* based on BDI architecture (Bratman 1987), norm life-cycle (Savarimuthu and Cranefield 2011; Frantz and Pigozzi 2018; Argente et al. 2020), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009). To evaluate *Noe*, we design a simulation experiment with various agent societies. We investigate how norms emerge and how emotions in normative agents influence social welfare.

To make the problem tractable, we apply one social norm in our simulation and simplify the emotional expression to reduce the complexity. Specifically, our *Noe* agents process emotions by appraising norm outcomes. For the emotion model, we adopt the OCC model of emotions (Ortony et al. 1988) in which we consider both valence and intensity of emotions and assume violation of norms yields negative emotions.

3.2 *Noe*

We now describe the architecture, norm formal model, and decision-making.

3.2.1 Architecture

Noe integrates the BDI architecture (Bratman 1987) with a normative model (Savarimuthu and Cranefield 2011; Frantz and Pigozzi 2018; Argente et al. 2020) and an emotional model (Alfonso Espinosa 2017; Marsella and Gratch 2009). A *Noe* agent assesses the environment, including other agents' explicit emotions, its cognitive mental states, and infer possible outcomes to make a decision. Figure 3.1 shows the three components of *Noe*.

The normative component of *Noe* includes the following processes:

- Identification: the agent recognize norms from its norm base based on its beliefs
- Instantiation: activate norms related to the agent
- Normative reasoning process: the reasoning process makes decisions based on the beliefs, current intention, self-emotions, other-emotions, active norms, and how the norm satisfaction or violation influences the world and itself. The *Noe* agents then update the intention based on the results of normative reasoning.
- Norm fulfillment process: check if a norm has been fulfilled or violated based on the selected action.

The BDI component includes the following parts:

- Beliefs: update beliefs based on perceptions.
- Desires: generate desires based on the beliefs
- Intention: the highest priority of desires to achieve based on the beliefs

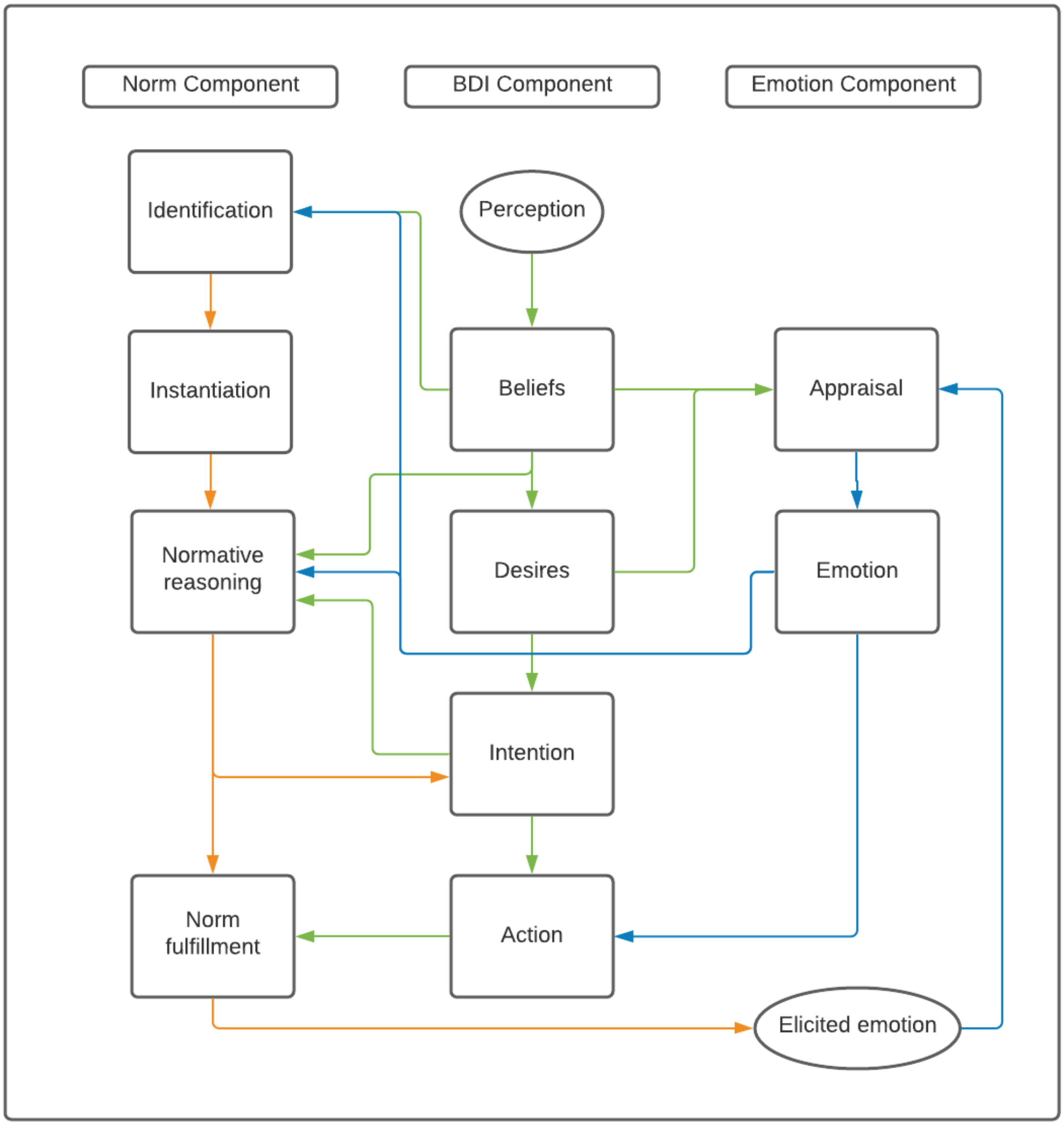


Figure 3.1: *Noe* Architecture: rectangles represent processes; circles represent data instances; arrows represent information flows.

- Action: select action based on the current intention, emotions, possible outcomes, and the evaluation of violating or complying with norms, if any

  The beliefs, desires, and intentions are mental states of *Noe* agents.

The emotional component includes the following processes:

- Appraisal: calculate the appraisal value based on the beliefs, desires, and the elicited emotion triggered by norm satisfaction or violation of a norm. In this work, we consider only the elicited emotion.

- Emotion: generate emotion based on the appraisal value (Marsella and Gratch 2009)

We consider both valence and intensity of emotions and assume violation of norms yields negative emotions.

Figure 3.2 illustrates the interactions between agents in our simulation scenario.

3.2.2 Norm formal model

Social norms describe the interactions between agents in a sociotechnical system. We adopt Singh's (Singh 2013) representation of social norms, where a social norm is formalized as $Norm(\text{subject}, \text{object}, \text{antecedent}, \text{consequent})$. In this representation, the subject and object are agents, and the antecedent and consequent are conditions under which the norm is activated or satisfied, respectively. This representation describes a norm activated by the subject towards the object when the antecedent holds, and the consequent indicates if the norm was satisfied or violated).

Following Singh (Singh 2013), we consider three types of norms in Noe.

- Commitment (C): the subject commits to the object to bring out the consequence if the antecedent holds. Consider Alice and Bob are queuing up in a grocery store. Alice and Bob commit to each other to keep social distance during the pandemic, represented as $C(Alice, Bob, during = pandemic, social\_distance)$.

- Prohibition (P): the object prohibits the subject from the consequence if the antecedent holds. Caleb, the grocery store manager, prohibits Bob from jumping the queue while lining up in that store, represented as $P(Bob, Caleb, when = lineup; at = grocery\_store, stay\_in\_queue)$.

- Sanction (S): same as commitment or prohibition, yet the consequence would be the sanctions. Sanctions could be positive, negative, or neutral reactions to any norm satisfaction or violation (Nardin et al. 2016). If Bob breaks the queue, he receives negative sanctions from Alice, represented as $S(Bob, Alice, jump, negative\_sanctions)$. Negative sanctions could be physical actions, e.g., scold, or emotional expression, e.g., expressions of disdain, scowl, or disgust.

To simulate the norm emergence and enforcement in human society, we include emotions into the decision-making process since, by nature, humans do not always act rationally in terms

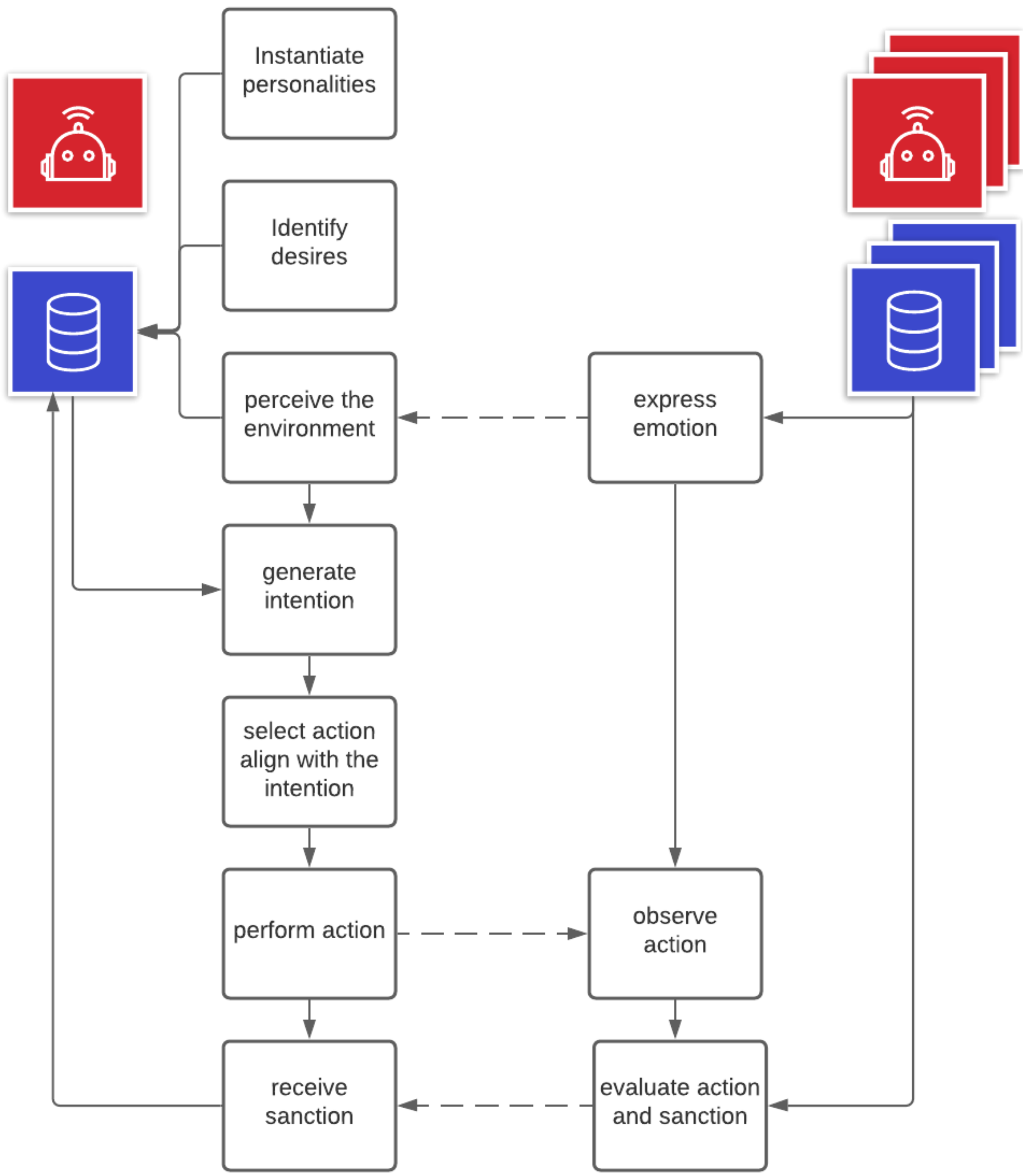


Figure 3.2: The interaction between *Noe* agents.

of utility theory. Here we formalize emotions with $E_i(target, intensity, decay)$ indicating agent a_i has emotion e toward the target with intensity and decay value. A target in this representation can be an object or an agent. An example of the prohibition case would be, Bob would not jump the queue if Alice is angry, represented as $P(Bob, Alice, Bob \succ Alice \wedge E_{Alice} = angry, stay)$.

We model the emotional response of agents with triggered emotions from norm satisfaction or violation (Argente et al. 2020). Here we represent the elicited emotions with $Elem_{name}(A_{expect}, A_{real}, Em_1, Em_2) | Em_1, Em_2 \in E; A_{expect}, A_{real} \in A$ where A is a set of actions. E is a set of emotions, and Em_1 and Em_2 are the emotions triggered by norm satisfaction and violation accordingly. If the A_{expect} is equal to the A_{real}, a norm has been fulfilled, and Em_1 was elicited. $Ap(beliefs, desires, Elem)$ represents the appraisal function.

3.2.3 Decision-Making

Schwarz (2000) addresses the influence of moods and emotions at decision making and discussed the interplay of emotion, cognition, and decision making. Specifically, the aspects include pre-decision affect, post-decision affect, anticipated affect, and memories of past affect. In our model, we include the pre-decision affect into the decision-making process. Specifically, people recall information from memories that match their current affect (Schwarz 2000). This recalled information affects people's decision-making process. Moreover, people in a sad mood particularly tend to overestimate adverse outcomes and events.

In our model, emotions serve as mental objects, which we represent with simple values where positive values indicate positive emotions and larger values indicate stronger emotions. We simplify from the OCC model (Ortony et al. 1988) and consider certain emotions in this work, i.e., fear, shame, reproach, and admiration. Specifically, we express emotions with numerical values where positive values indicate positive emotions and the value of emotions indicates its intensity. We ignore mood since mood is hard to detect. We only consider short-term emotions. *Noe* agents' appraisal function considers norm satisfaction only. The agents are aware of other agents' expressed emotions in the same place, their own desires and beliefs, and available actions to achieve their goals. In this work, we assume that agents express true and honest emotions and perceive expressed emotions. In other words, felt emotions are equal to expressed emotions. Another assumption is that emotions are consistent with the notions of rational behavior.

Algorithm 1 displays the decision loop of our model. At the beginning of the simulation, all agents are initialized with certain desires, and during the run, an intention would be generated by prioritizing desires with the agent's beliefs. When choosing the next move with line 5 in Algorithm 1, the agent chooses the one with maximum utility from all available actions. Algorithm 2 details the action selection. The decision takes the agent's beliefs, current intention, and possible consequences into accounts. While norms are activated with the beliefs, the agent would further consider emotions and cost and possible consequences with norms at line 9 in

Algorithm 2. For instance, if people violate some social norms, they may be isolated from society. Regarding the influence of emotions, people may overestimate the negative outcomes when they are in the negative emotion and tend to comply with the norms.

Algorithm 1: Decision loop of a *Noe* agent

```
/ * initialize one agent with its desires D * /
while True do
    observe the environment and update beliefs B;
    generate intention I based on B and D;
    a = ActionSelection(B, I, D, E);
    execute the selected action a;
    if action a fulfills a norm then
        elicit positive or neutral emotions E from agent itself and observer agents;
        # other-directed emotions
        E as sanctions to others;
    else
        elicit negative or neutral emotions E from agent itself and observer agents;
        # other-directed emotions
        E as sanctions to others;
    end
    # self-directed emotions
    E as sanctions to itself;
end
```

3.3 Evaluation

We evaluate *Noe* via a line-up environment where agents form queues to receive service. We detail the environment in Section 3.3.1.

3.3.1 Line-up Environment

Figure 3.3 shows the line-up environment. We build this line-up environment using Mesa (Masad and Kazil 2015), a Python-based framework for building, analyzing, and visualizing agent-based models.

The line-up environment includes two shared locations—home and grocery stores. The

Algorithm 2: Action selection

```
/ * choose one action that maximizes utility * /
Input: beliefs, intention, desires, and emotions
Output: Action a
Function Action Selection:
    for available actions do
        N = activated norms with beliefs;
        if N = ∅ then
            selected_action = max_utility(beliefs, intention, possible_action)
        else
            for N do
                # possible result on self and others
                possible_result = norm_reasoning(N, beliefs, intention, possible_action)
                  * amplifier(emotions)
                selected_action = max_utility(possible_result)
            end
        end
    end
    return selected_action
return
```

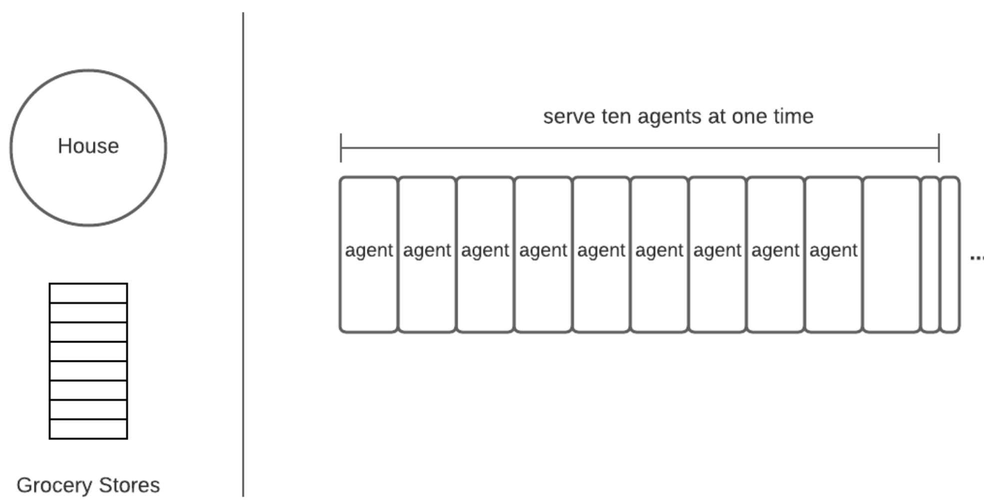


Figure 3.3: Simulation details. The left part: accessible locations; The right part: a queue for agents to line up to enter the grocery stores. Only N agents get services at one time.

agents move between home and grocery stores to get food. We consider one social norm in the line-up environment — to enter the grocery store, agents are expected to line-up. To simulate real human reactions to norm violations, we refer to a social psychology experiment (Milgram et al. 1986). In the line-up environment, we model defensive reactions of people in the queue as negative emotions toward those who jump the queue by barging in ahead of someone already in the queue. People show positive emotions toward those who stay in the queue.

We initialize the agents with the following parameter values:

- Health (Integer value from 0–100): When the health value reaches zero, the agent is marked as *deceased* and unable to act. The health value decreases by 1 unit at each step.
- Deceased (Boolean: True or False): the agent's state; marked as True when an agent runs out of health.
- Emotion (Integer value): simplified with numerical values where positive values indicate positive emotions and negative indicate negative emotions. The emotions come along with a duration. Default at 0.
- Number of food packets owned (Integer value from 0–15): once obtained food from the stores, agents would be able to restore its health value via consuming food anywhere.
- Food expiration day (Integer value from 0–15): once the agent gets food packets, we update the expiration day with 15. The expiration day decreases by 1 unit at each step. Food expires once the expiration day reaches 0. Default at 0.
- Beliefs: the perceived and processed information from the world, including location, other agents' emotions.
- Desires: desired states, including *have food* and *wandering*.
- Intention: the highest priority of desires to achieve at a specific time. When the agent's health is lower than the threshold, 80% of the health, the agent sets its intention as *get food*; otherwise, the agent sets its intention as *wandering*.

When an agent runs low on stock, it has a higher probability of moving to a grocery store. The grocery store can provide food packets to eight agents in one time step. While waiting in line to get food, the agent could either stay in the line or jump ahead in the line to get food with less time. Jumping the line may increase other agents' delay in getting food packets. Those who witness the violation would then cast negative emotions, further interpreted as anger or disdain, triggered by that behavior. To simplify the simulation, we presume the anticipated affects (Schwarz 2000) with: (1) receiving negative emotions triggers negative self-directed emotions such as shame and guilt; (2) complying with norms leads to positive or neutral emotions; (3) violating norms leads to negative or neutral emotions. The intensity of emotions triggered each time is fixed but can accumulate. Each triggered emotion lasts 2 steps. At each step, the duration and intensity of emotion decrease by 1 as decay. The values of emotions can add up. A simple assumption here is that people in a bad mood would trigger stronger emotions in response to a non-ideal state. Note that at the beginning of the simulation, we initialize the agent society with health in normal distribution to avoid all agents having the same intention at the same time.

3.3.2 Agent Types

To answer our research question and evaluate *Noe*, in addition to *Noe*, we define three types of agent societies as baselines. We describe the agents societies below:

Obedient society. Agents in an obedient society always follow norms.

Anarchy society. Agents in an anarchy society jump lines when they cannot get food.

Sanctioning society. Agents in the sanctioning society jump lines considering the previous experience of satisfying or violating a norm. Agents sanction positively or negatively based on norm satisfaction or violations directly and comply with enforced norms.

***Noe* society.** Agents in the *Noe* society jump lines considering the previous experiences of satisfying or violating a norm, current emotional state of the other agents, current self emotional state, and estimated outcome of satisfying or violating a norm. *Noe* agents who observe norm satisfaction or violations would appraise the norm outcomes and trigger emotions to sanction the actor agent.

Table 3.1 summarizes the characteristics of the agents in the four societies.

3.3.3 Hypotheses and Metrics

To address our research question $\text{RQ}_{\text{emotion}}$ on emotions and norm emergence, we propose two hypotheses:

$\mathbf{H_1}$ (Norm satisfaction): Norm satisfaction in *Noe* agent society is higher compared to the baseline agent societies.

Table 3.1: Characteristics of *Noe* and the baseline agent societies.

Agent Type	Violation allowed	Sanctioning	Emotions involved
Obedient society	✘	✘	✘
Anarchy society	✔	✘	✘
Sanctioning society	✔	✔	✘
Noe society	✔	✔	✔

H_2 (Social welfare): *Noe* agent society yields a better social welfare compared to the baseline agent societies.

H_3 (Social experience): *Noe* agent society yields a better social experience compared to the baseline agent societies.

To evaluate H_1 on norm satisfaction, we compute one metric:

M_1 Cohesion: Percentage of norm satisfaction

To evaluate H_2 on social welfare, we compute three metrics:

M_2 Cumulative number of agents deceased

To evaluate H_3 on social experience, we compute three metrics:

M_3 Average waiting time of agents in the queues

M_4 Average health of the agents

To test the statistical significance of H_1, H_2, and H_3, we conduct the independent t-test and measure effect size with Glass's Δ for unrelated societies (Grissom and Kim 2012; Glass 1976). We adopt Cohen's (Cohen 1988) descriptors to interpret effect size where above 0.2, 0.5, 0.8 indicate small, medium, and large.

3.3.4 Experimental Setup

We run each simulation with 400 agents and queue size 80 for 3,000 steps. To reduce the simulation time, we choose a relatively small number of agents. Our results are stable for a larger number of agents. The simulation stabilizes at about 1,500 steps of our simulation, but we keep extended simulation steps to have more promising results.

We present the results with an average of a running window of 100 steps. We choose this size of running window to show the temporal behavior change in a small sequence of time. With a larger size, the running window may alleviate the behavior change. To minimize deviation from coincidence, we run each simulation with 10 iterations and compute the mean values.

3.3.5 Experimental Results

In this section, we describe the simulation results comparing the three baselines and *Noe* agents. Table 3.2 summarizes the experiment results and lists the value of Glass's Δ and p-values from the independent t-test.

According to Table 3.2, we see that *Noe* generate better cohesion and social welfare than baselines ($p < 0.01$; Glass's $\Delta > 0.8$). The null hypotheses corresponding to H_1 and H_2 are rejected. Note that we do not consider the cohesion metric for the obedient agent society here since agents in the obedient society are always compliant. However, *Noe* also yields the worst social experience where low waiting time and high health are desirable states ($p < 0.01$; Glass's $\Delta > 0.8$). The null hypothesis corresponding to H_3 indicates no significant difference.

Table 3.2: Comparing *Noe* agent society with baseline agent societies on various metrics and their statistical analysis with Glass' Δ and p-value.

		Obedient	Anarchy	Sanctioning	*Noe*
$H_{\text{Norm Satisfaction}}$	M_{Cohesion}	1	0.22	0.88	0.99
	p-value	0.32	< 0.01	< 0.01	–
	Δ	0.19	102.43	13.67	–
$H_{\text{Social welfare}}$	M_{Deceased}	55.30	81.60	169.30	**54.00**
	p-value	< 0.01	< 0.01	< 0.01	–
	Δ	0.65	3.1	15.53	–
$H_{\text{Social experience}}$	M_{Waiting}	8.95	5.45	2.55	8.95
	p-value	0.98	< 0.01	< 0.01	–
	Δ	0.01	40.82	76.68	–
	M_{Health}	79.27	79.50	86.26	79.00
	p-value	0.52	0.46	8.45	–
	Δ	0.18	0.21	3.34	–

H_1 Norm Satisfaction

Figure 3.4 displays the cohesion, the percentage of norm satisfaction, in the baseline agent societies and the *Noe* agent society. We find that the percentage of norm satisfaction in the *Noe* agent society, average at 99% and p-value < 0.01, is constantly higher than the sanctioning agent society, average at 88% and p-value < 0.01 and Glass's $\Delta > 0.8$. The sanctioning agent society learns to comply with the norm as time goes by. The *Noe* agent society does sanction as well. Yet, considering emotions and the possible outcome makes *Noe* agent society enforce the norm faster than the sanctioning agent society. Specifically, *Noe* agent society enforces the norm at about 100 steps while sanctioning agent society at 1500 steps.

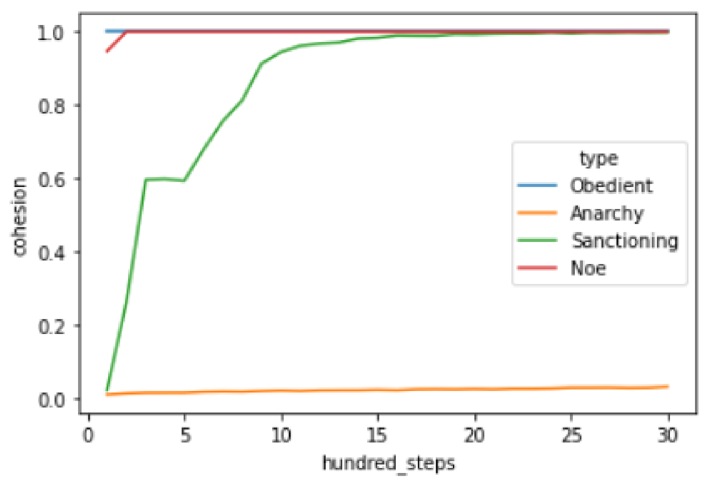


Figure 3.4: Simulation result: average cohesion. Comparing average cohesion (M_1) yielded by *Noe* and baseline agent societies.

H_2 Social Welfare

Figure 3.5 compares the average number of deceased in the obedient, anarchy, sanctioning, and *Noe* agent societies. Refer to Figure 3.4, sanctioning agent society soon learns the norm with positive and negative sanctioning from norm satisfaction and violation. However, the agents in that society do not consider the outcome of norm satisfaction to cause compliant agents to die in the queue. When the number of deceased reaches the threshold, the simulation stabilizes. Therefore, no more agent from the sanctioning agent society dies after the threshold. On the contrary, *Noe* agent society sanctions and considers possible outcomes of norm satisfaction and violation, therefore learning the norm and avoiding unacceptable consequences.

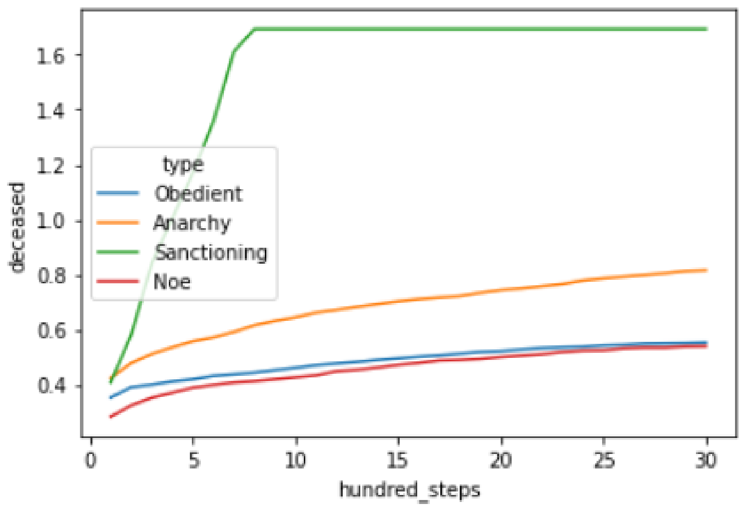


Figure 3.5: Simulation result: average number of deceased. Comparing average number of deceased (M_2) in *Noe* and baseline agent societies.

H_3 Social Experience

Figure 3.6 compares the average duration the agents spend in a queue at the grocery store in the obedient, anarchy, sanctioning, and *Noe* agent societies. The *Noe* agent society learns the norm fast and remains the same waiting time in the queue. However, some agents in the sanctioning agent society take advantage of those who learn norms faster than themselves. Therefore, many agents die during the learning process, and the simulation stabilizes. In Figure 3.6, the obedient agent society shares the same trend with *Noe* agent society since emotions enforce the line-up norm.

Figure 3.7 compares the average health of the agents in the obedient, anarchy, sanctioning, and *Noe* agent societies. The sanctioning agent society performs better, average at 86.26, since most agents die in the learning process. The rest of the agents then be able to remain in high health.

Combining the results for H_1 and H_2 and H_3, we note that while sanctioning enforces norms, a combination of sanctioning and emotions enforce norms better. Specifically, having emotions as amplifiers of outcomes yield higher norm satisfaction compared to our baselines. The results also indicate that, first, sanctioning agents that consider only norm violation or norm satisfaction

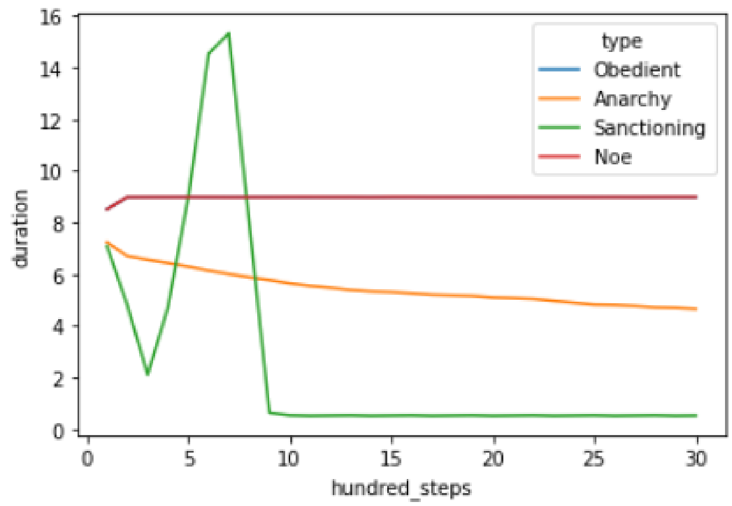


Figure 3.6: Simulation result: average waiting time of agents in queues. Comparing average waiting time (M_3) in *Noe* and baseline agent societies.

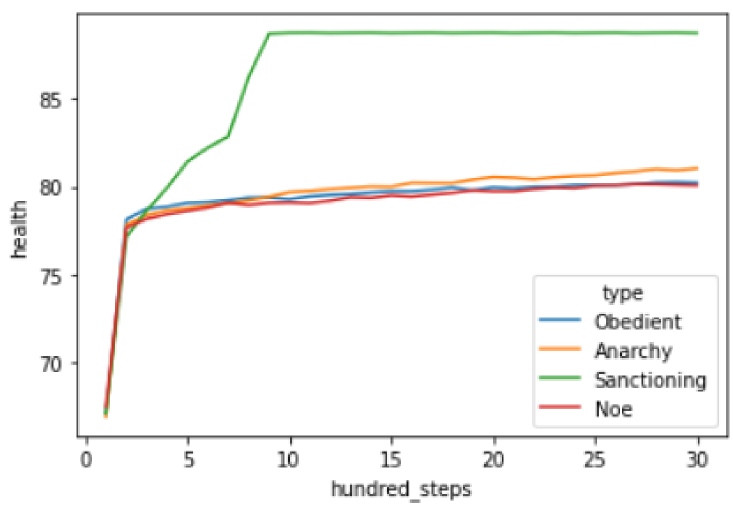


Figure 3.7: Simulation result: average health value. Comparing average health value (M_4) in *Noe* and baseline agent societies.

may bring out worse social welfare compared to *Noe* that considers both norms and their consequences. Second, although *Noe* agents remain relatively high waiting time in the queues, the number of deceased is lower than the baselines. Note that the sudden drop of deceased number or increase of health value for sanctioning agents resulted from the stabilization of that society. Third, *Noe* agents stay in positive emotions during the simulation while sanctioning agents start from negative emotions and finally achieve the expected behaviors.

3.4 Discussion and Conclusion

We present an agent architecture inspired by norm life-cycle (Argente et al. 2020), BDI architecture (Bratman 1987), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009) to investigate the interactions of norms and emotions. We evaluate the proposed architecture via simulation experiments in an environment where agents queue up to receive service. In our simulations, we consider two characteristics of an agent society: sanctioning and emotions that participate in action selection and arise from evaluating selected action.

As a future extension of current work, we plan to differentiate emotions in *Noe* instead of general emotions to provide more information. We also consider including a mix of personality in future research to have different appraisal results. In this work, *Noe* agents are assumed to express true and honest emotions, yet in an adversarial context, emotions can also serve as tools to influence, persuade or deceive others.

CHAPTER

4

EXPRESSED EMOTION AS INFORMATION

4.1 Introduction

The widespread COVID-19 pandemic has caused substantial economic costs and human suffering (D'Orazio et al. 2021). During a pandemic, humans need a reliable model to understand each policy's possible costs and consequences to minimize the expenses (Dignum et al. 2020). Although many epidemics and social science models perform well on predicting the varied outcomes of a pandemic, these models ignore the influence of individual behaviors (Allen 1994). Specifically, these models obtain the overall numbers from calculations with probabilities and lack the understanding between the spread of pandemics and individual behaviors.

In human society, social norms regulate behaviors (Savarimuthu and Cranefield 2011; Singh 2013), but humans have the capability to deviate from norms in certain contexts. For instance, people have to deviate from their regular social norms, e.g., socializing with friends, and adopt lockdown regulations during a pandemic (Lau et al. 2020). Sanctions are reactions to the compliance or violation of social norms, which can be positive, negative, or neutral (Nardin et al. 2016). Norms either are defined in a top-down manner or emerge in a bottom-up manner (Savarimuthu and Cranefield 2011). In the bottom-up approach, agents learn common behaviors from interacting with others (Morris-Martin et al. 2019) while in the top-down approach, agents are given norms from a legalistic view, e.g., legal norms.

4.1.1 Motivating Scenarios

We now discuss how norms exist in our daily life and what motivates *Hermione*.

Example 1 ***Punishment.*** *Bella becomes infected with COVID-19 after having direct contact with an infected individual. Bella becomes symptomatic and infectious after the latent period, within which an individual is infected but not infectious. Asher goes out for grocery shopping and meets Bella in a grocery store. During the interaction, Asher notices Bella's symptoms. To better handle some extreme conditions, e.g., a pandemic, Taiwan's one communicable disease control act, similar to the executive order in the United States, grants the Center for Disease Control (CDC) the power to coordinate local authorities. According to COVID-19 guidelines in Taiwan, individuals who have direct contact with infected people or have COVID-19 symptoms should be tested, and quarantine at home (Nussbaumer-Streit et al. 2020; Wang et al. 2020). Anyone who violates this policy and is reported to the CDC would face fines from the authorities and be forced into home quarantine. When Asher perceives Bella's symptoms and that Bella is not self-quarantining, Asher may negatively sanction Bella via reporting her to the local authorities. Bella would then be forced into home quarantine and face fines for her violation of COVID-19 guidelines.*

Although such regulations perform well on disease control, a massive amount of human resources, well-structured surveillance systems, and related technology are required. In other words, such regulations pose an extravagant cost to society. However, real-world sanctions are often more subtle than the above. Specifically, the sanctions could be verbal warnings, emotions, or changes of reputation that might change one's behavior (Nardin et al. 2016; Bourgais et al. 2019). Typical punishing strategies in the real world naturally include normative information and material punishment (Andrighetto et al. 2013). Emotions, the responses to internal or external events or objects, influences the decision-making process and provide extra information in communication (Keltner and Haidt 1999; Schwarz 2000). Tzeng et al. (Tzeng et al. 2021) considered their emotions as sanctioning approach and the critical factor that changes how an agent evaluates current states. Humans' decision-making over norms reflects their preferences and attitudes (Ajmeri et al. 2018). In some scenarios, humans do not always act rationally in terms of utility theory. For instance, Asher drives along Research Road where no surveillance systems and no one else is in sight. Since Asher cannot wait to go home, he intends to ignore the speed limit. However, the well-known truth is that state polices patrol this area more often at night. Considering his fear of being caught and issued fines, Asher drives within the speed limit.

Apart from being a way to sanction or being a factor to explain some human behaviors, expressed emotions as information enable inference of mental states that are otherwise not observable (Wu et al. 2018a; Wu and Schulz 2020). Compared to explicit expression, emotions show subtle normative information over behaviors. With explicit messages, humans gain direct and indirect normative information. An example of indirect messages is folklore in many countries. For instance, Santa Claus will give good kids toys. Children who believe in this story may then learn to behave themselves and expect to get rewards. The example of direct messages would be Example 2.

Example 2 ***Normative Message.*** *Cecilia notices Bella's suspicious symptoms and begins to worry about her safety. Considering quarantine at home isolates potentially infected people and reduces healthy individuals' risk (Wang et al. 2020), Cecilia hopes Bella will self-quarantine. Since Cecilia feels a threat to her life, she warns Bella to stay home while she has symptoms or she will be reported to the local authorities. Upon receiving such information, Bella fears the possible penalty and changes her behavior accordingly. Other people who observe this event also learn the causal link between COVID-19 symptoms and being forced into home quarantine.*

While messages provide clear normative information, emotions give subtle normative information for our behavior as well. Upon receiving negative emotions after some actions, we can infer that our behaviors do not fit into others' expectations.

Example 3 ***Emotion As Information.*** *David notices Bella's suspicious symptoms and expresses little negative emotions with facial expressions, e.g., anger. Upon perceiving the emotion, Bella might infer the emotion was for her violation of self-quarantine and believes that some potential punishments may happen. Bella then changes her behavior. Other people who observe this may make the same inference and learn the causality.*

4.1.2 Research Contributions

In multiagent systems, modeling and reasoning about emotions and other affective characteristics in an agent become important. A mass of researches showed a connection between human emotions and adaptive decision-making. In this work, we argue that normative information from both messages and emotions provide hints of future outcomes and therefore promote cooperation.

We apply agent-based models (ABM), which provide granularity, to simulate realistic disease transmission. To reduce human intervention and efforts, we adopt reinforcement learning and show that reinforcement learning has the potential to model norms and emotions. One challenge of epidemic research is that it is difficult to capture all emergent human behavior due to limited data. However, we focus on how emotion influences norm emergence and the robustness of norms in this work.

We investigate the following research questions.

RQ_{RL}. How does reinforcement learning accommodate reasoning about cognitive constructs, emotions, and norms?

$\text{RQ}_{\text{information}}$. How does providing indirect information, e.g., emotion as information, influence norm emergence?

To address $\text{RQ}_{\text{information}}$, we define two expressions: explicit normative explanation (Andrighetto et al. 2013) and emotion as information. Consider Example 2, with explicit normative explanation, Cecilia would state normative information and possible sanctions. With emotion as information in Example 3, Bella receives the expressed emotions from David and makes inferences from the emotions and the quarantine at home regulation. We apply belief reward shaping (Marom and Rosman 2018), a reward augmentation framework that considers rewards from the environment and also from beliefs, in our simulation. To answer $\text{RQ}_{\text{information}}$, we design a simulation with reinforcement learning (RL) and *Hermione*. Our proposed framework *Hermione* combines a cognitive architecture (simplified from the Belief-Desire-Intention (BDI) architecture (Rao and Georgeff 1991)), norm life-cycle (Argente et al. 2020; Frantz and Pigozzi 2018; Savarimuthu and Cranefield 2011), and emotion life-cycle (Alfonso Espinosa 2017; Marsella and Gratch 2009).

To make our evaluation of $\text{RQ}_{\text{information}}$ tractable, we apply one social norm in our simulation and simplify our treatment of emotions to primary appraisal (Lazarus 1991). Specifically, our *Hermione* agents process emotions by appraising norm outcomes along with the likelihood of reaching their's desires. For the emotion model, we adopt the model of emotions proposed by Ortony, Clore, and Collins (OCC) (Ortony et al. 1988), in which we consider the valence of emotions and assume violation of norms yields negative emotions.

Summary of findings First, we integrate reinforcement learning with cognitive constructs, emotions, and norms. Second, we show that, akin to normative messages, emotional expressions provide normative information and promote cooperation.

4.2 Background

We now provide preliminary materials necessary to understand our contribution.

4.2.1 Q-Learning

Q-Learning (Watkins and Dayan 1992) is a model-free reinforcement learning algorithm that learns from trial and error with given rewards or penalties. The reward function in Q-Learning defines the desired or undesired states. Q-Learning algorithm computes the action-state value $Q(s,a)$ (Q value), which indicates the expected and cumulative rewards for each state and action. By approximating the value of an action for a given state, the Q-Learning algorithm finds the optimal policy. The Q function that computes Q values with the weighted average of the old value and the new information::

$$Q'(s_t,a_t) = Q(s_t,a_t) + \alpha * (r_t + \gamma \max_{a'} Q(s_{t+1},a) - Q(s_t,a_t)) \tag{4.1}$$

where $Q'(s_t,a_t)$ represents the updated cumulative value of $Q(s_t,a_t)$ after performing action a at time t. In this equation, α indicates the learning rate, and γ defines the reward discount rate.

In multiagent reinforcement learning (MARL), agents must learn how to interact with other agents in a shared environment, including more complexity and uncertainty posed by other agents. The simplest form of MARL is independent reinforcement learning (InRL), where the learning agent treats other agents as part of its nonstationary environment and makes decisions based on its local observations.

4.2.2 Belief Reward Shaping

Reward shaping (Ng et al. 1999) provides additional "shaping" reward from deterministic reward function: $F : S \times A \times S \rightarrow \mathbb{R}$. With reward shaping, the reward function of the transformed Markov Decision Process (MDP) becomes $R' = R + F$ where R is the actual reward and F is the belief reward. Here, $F^\tau = \hat{p}^\tau(r|h) : h \in H^\tau$ where τ denotes a transition and H^τ denotes the the hypothesis space for transition τ, and $\hat{p}^\tau$ is a hypothesis for the true environment reward.

4.3 *Hermione*

We now describe *Hermione*'s architecture, formal norm model, and decision-making.

4.3.1 Architecture

Hermione instantiates a cognitive architecture (Rao and Georgeff 1991) based on a normative model (Argente et al. 2020; Frantz and Pigozzi 2018; Savarimuthu and Cranefield 2011) and an emotional model (Alfonso Espinosa 2017; Marsella and Gratch 2009). Figure 4.1 shows the three components of *Hermione*. A *Hermione* agent makes decisions with its cognitive states, and incorporates normative reasoning in the decision-making process. The cognitive component of *Hermione* describes an agent's beliefs, desires, and intentions. A *Hermione* agent forms beliefs based on its perceptions, including information from the environment, information from other agents, and its available actions. Considering Example 2, Bella believes that she will be reported to local authorities if he does not self-quarantine upon receiving Cecilia's message. Since environments are not fully observable in most cases, beliefs are not necessarily the true states. In Example 3, upon noticing David's emotional expression, Bella believes that David disapproves of her self-quarantine violation. However, the actual situation may be that David is angry because of his own performance in class.

Hermione's emotion component specifies the process of emotion generation. The appraisal function first evaluates an agent's beliefs, desires, and the emotion-events triggered by norm satisfaction or violation of a norm. After the appraisal function generates emotions based on the evaluation, the emotion process updates the agent's affective states. We consider the valence of emotions and assume that violation of norms yields negative emotions.

The normative component of *Hermione* describes the life-cycle of norms. A *Hermione* agent recognizes norms based on its beliefs from its knowledge base in the identification process and activates those norms related to itself in the instantiation process. A *Hermione* agent considers norm compliance or violation for the decision module in the normative reasoning process. After executing the chosen action, the norm fulfillment process checks if a norm has been fulfilled or violated based on what action was performed. The compliance and violation of norms then trigger emotions.

Figure 4.2 illustrates the interactions between agents in our simulation scenario. Agents can perceive their environment, including other agents' actions. In Example 1, when Asher finds Bella at the grocery store, he may sanction Bella after the decision-making process.

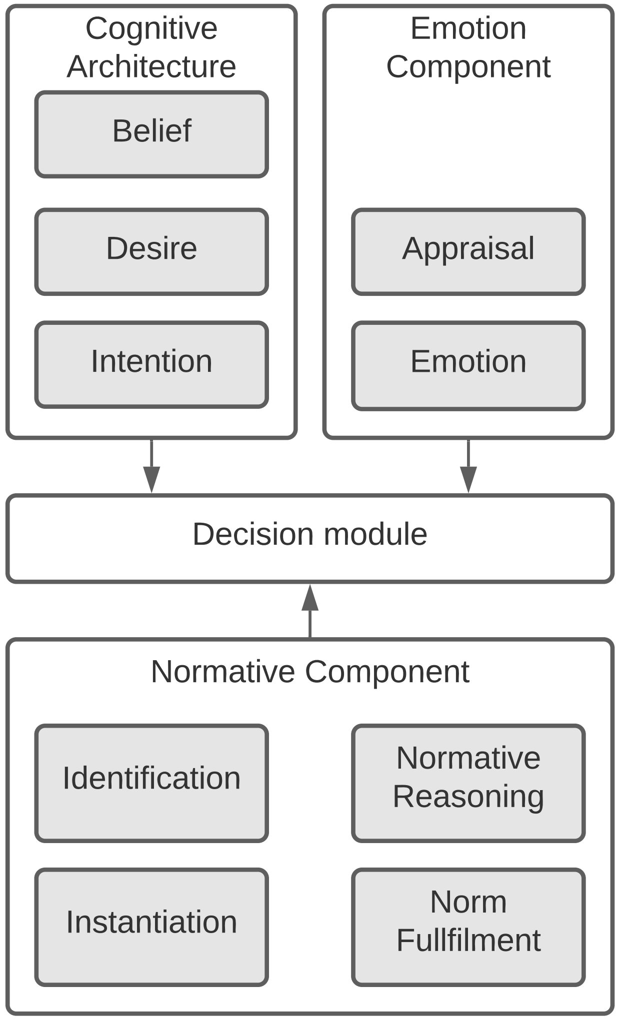


Figure 4.1: *Hermione* architecture, representing and reasoning over beliefs, desires, intentions, emotions, and norms.

4.3.2 Formal Norm Model

Social norms describe interactions between agents in a sociotechnical system. We adopt Singh's (Singh 2013) representation of social norms, where a social norm is formalized as $Norm(\text{subject}, \text{object}, \text{antecedent}, \text{consequent})$. In this representation, the subject and object are agents, and the antecedent and consequent are conditions under which the norm is activated or satisfied, respectively. This representation describes a norm that focuses on the subject and arises on the object while the antecedent holds. The consequent indicates the conditions under

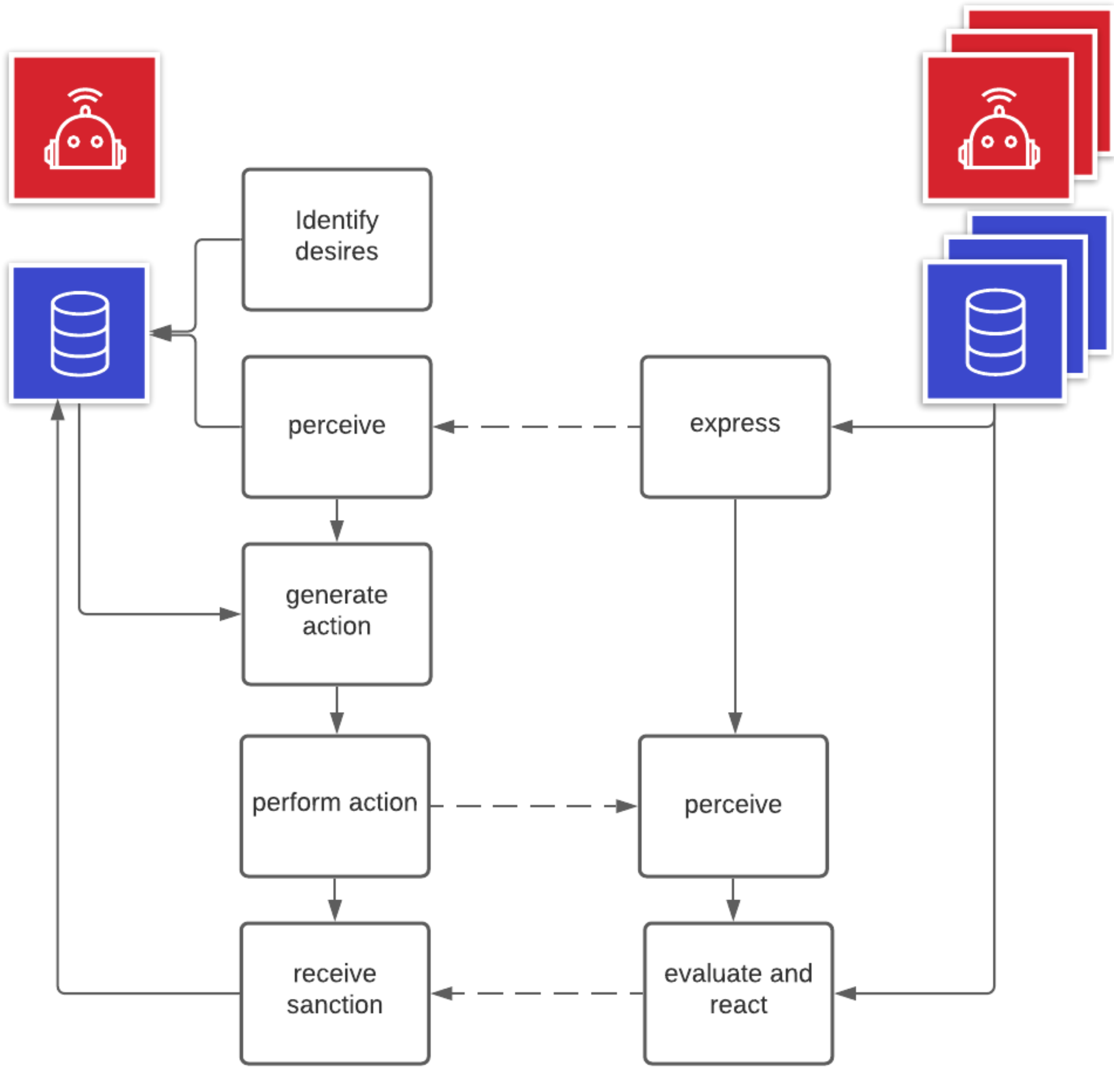


Figure 4.2: The interaction between *Hermione* agents.

which the norm is satisfied. Table 4.1 describes the syntax of *Hermione*. In the syntax, $\mathscr{A}$ represents agent id and Expr represents composite states. Emotion represents emotion valence that indicating intrinsic attractiveness and aversiveness (Frijda 1986).

Following Singh (Singh 2013), we consider three types of norms in *Hermione*.

- Commitment (C): the subject commits to the object to bring out the consequent if the antecedent holds. Suppose that during the outbreak of COVID-19, the government were to announce a self-quarantine regulation. That is, each individual who tests positive for

COVID-19 must self-quarantine for fourteen days. The government may issue fines to those who violate this policy and send them to self-quarantine. To improve the control and prevention of the COVID-19, the government recommends individuals report any violation. Suppose Asher commits to the government to stay self-quarantine, represented as $C(\text{Asher}, \text{government}, \text{infected}, \text{stay home})$.

- Prohibition (P): the object prohibits the subject from the consequent if the antecedent holds. The government prohibits Bella from going out while Bella is infected, represented as $P(\text{Bella}, \text{government}, \text{infected}, \text{go out of home})$.

- Sanction (S): the object sanctions the subject by bringing out the consequent if the antecedent holds. Sanctions could be positive, negative, or neutral reactions to any norm satisfaction or violation (Nardin et al. 2016). To better handle the highly infectious COVID-19, some governments introduce enforceable rules. For instance, people who violate the self-quarantine policy or refuse to cooperate with orders in Taiwan or Japan could face fines. Following the previous example, if Bella does not isolate herself from others after being infected, she receives negative sanctions from Asher, represented as $S(\text{Bella}, \text{Asher}, \text{infected}; \text{go out of home}, \text{negative sanctions})$. Negative sanctions could be legal actions, e.g., penalize with fines, ostracism, disseminating criticism, or emotional expressions, e.g., facial expressions of disdain, scowl, or disgust.

We incorporate emotions into the decision-making module of *Hermione* and simulate norm emergence and enforcement in an agent society. The OCC model (Ortony et al. 1988) describes a hierarchy of emotion types, which contains three branches: (1) emotions regarding consequences of events (e.g., joy, relief, and distress), (2) actions of agents (e.g., pride and admiration), and (3) attractions of objects (e.g., liking and disliking). Here, we formalize emotions with $E_i(\mathscr{A}) = valence$ indicating agent a_i has emotion with valence as value towards the agent $\mathscr{A}$. Continuing with Example 3. Bella would not hang out in public space after being infected if David is angry, represented as $P(\text{Bella}, \text{David}, \text{infected}; E_{David}(Bella) = negative, \text{go out of home})$.

We model the emotional response of agents with emotion-events triggered from norm satisfaction or violation (Argente et al. 2020). Here, we represent the appraisal function of agent i with $Appr_i(\mathscr{A}, \text{Norm}_i, \text{Desires}_i, \text{Belief}_i)$ where Norm_i denote agent i's norm satisfaction or violation. Belief_i represents agent i's current beliefs and Desires_i represents agent i's desires. If Bella feels guilty for violating the self-quarantine regulation, we present with $Appr_{Bella}(\text{Bella}, \text{Norm}_i = False, \text{shopping}, \text{go out of home}) = negative$. From Asher's perspective, Bella's behavior is blameworthy. $Appr_{Asher}(\text{Bella}, \text{Norm}_i = False, \text{shopping}, \text{go out of home}) = negative$.

4.3.3 Information Expression

The normative information includes the causal link between preconditions and the consequences, which we present with $M(\mathscr{A}, \mathscr{A}, Expr, Expr)$. The first Expr represents the cause, and the second Expr represents the effect. The first and second $\mathscr{A}$ indicate the message sender and receiver, respectively.

For instance, if Bella stays in a grocery store with COVID-19 symptoms displayed, Asher may warn her first instead of giving costly hard sanctioning immediately. Precisely, hard sanctioning means forced quarantine in our scenarios. Asher would send:
$M(\text{Asher}, \text{Bella}, \text{infected}; \text{go out of home}, \text{negative sanction})$

In contrast to direct messages, emotion serves as an indirect normative information source. When Asher reports Bella to the local authorities and shows negative emotions, Bella may infer that her behavior triggers negative emotions and further associates negative emotions with the quarantine at home regulation.

Table 4.1: *Hermione* Syntax.

Norm	$\longrightarrow$	Commitment | Prohibition | Sanction
Norm	$\longrightarrow$	N($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)
Commitment	$\longrightarrow$	C($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)
Prohibition	$\longrightarrow$	P($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)
Sanction	$\longrightarrow$	S($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)
Message	$\longrightarrow$	M($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)
Belief	$\longrightarrow$	Expr
Desires	$\longrightarrow$	Expr
Appraisal	$\longrightarrow$	Appr($\mathscr{A}$, Norm, Desires, Beliefs)
Emotion	$\longrightarrow$	E($\mathscr{A}$)
Emotion Valence	$\longrightarrow$	positive | negative | neutral
Expr	$\longrightarrow$	true $\|\phi\|\neg$ Expr | Emotion | Expr $\wedge$ Expr

4.3.4 Decision-Making

Schwarz (Schwarz 2000) addresses the influence of moods and emotions in decision-making and highlights the close interplay between feeling and thinking in judgment and decision-making. Specifically, Schwarz discusses pre-decision affect, post-decision affect, anticipated affect, and memories of past affect. We include post-decision affect into our model. Briefly, negative outcomes may lead to regret and disappointment (Schwarz 2000).

We simplify from the OCC model (Ortony et al. 1988) and consider emotion valence without arousal in our model, Furthermore, emotions serve as mental objects in *Hermione*. Since mood lasts long and is hard to identify, we exclude mood from our model and only consider emotions. *Hermione* agents' appraisal function considers normative satisfaction and violation, beliefs, and desires. *Hermione* agents are aware of the information from other agents in the same place, the information from the environment, and their operations. Specifically, *Hermione* agents can only communicate with nearby agents. We assume that emotions are based on or are consistent with notions of rational behavior. That is, no emotions are based on the need to influence or deceive others. We further assume that agents can perceive each other's expressed emotions, and all expressed emotions are true and honest.

We describe the general causal relationships in *Hermione* and how they are justified in this section. The justification applies to commitment, prohibition, authorization, prohibition, and sanction in (Singh 2013). Take a prohibition scenario for instance. Suppose Asher (A) and Bella (B) interact with each other within a pandemic context. Figure 4.3 shows B's current decision to satisfy its prohibition towards A ($\mathsf{N}_{\mathsf{B,A}}$, abbreviated to N($\mathscr{A}$, $\mathscr{A}$, Expr, Expr)). The norm in this scenario is a prohibition. First, the notions include:

- Norm $\mathsf{N}_{\mathsf{A,B}}$. A Norm $\mathsf{N}_{\mathsf{A,B}}$ defines the relationship between an individual A on whom the norm is focused and an individual B with reference to whom the norm arises from bringing about the consequent when the antecedent holds (Singh 2013). The outcome of a norm can either be violated or satisfied when the consequent holds or not respectively.

- Sanction $\mathsf{S}_{\mathsf{A,B}}$. A sanction $\mathsf{S}_{\mathsf{A,B}}$,equivalent to S($\mathscr{A}$, $\mathscr{A}$, Expr, Expr), means that in a given context, an individual B would sanction an individual A through bringing about the consequent when the antecedent holds (Singh 2013). The outcome of a sanction has a Boolean value.

- Desire D_{A}. A desire D_{A} is a condition that an individual A wants to achieve. The outcome of a desire has a binary value, achieved or failed.

- Emotion $\mathsf{E}_{\mathsf{A,B}}$. An emotion $\mathsf{E}_{\mathsf{A,B}}$ describes a mental state an individual A has towards individual B. An emotion $\mathsf{E}_{\mathsf{A,B}}$ has three values: negative, neutral, or positive.

Here we further explain the details:

- $\mathsf{D}_{\mathsf{A}} \to \mathsf{E}_{\mathsf{A,B}}$ represents that the past outcomes of A's desires influence A's current emotions.

- $\mathsf{N}_{\mathsf{B,A}} \to \mathsf{E}_{\mathsf{A,B}}$ represents that the outcomes of B's past norms influence A's current emotions.

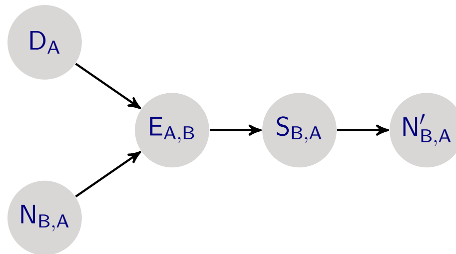


Figure 4.3: Causality in decision making: B's decision to satisfy its norm towards A. A denotes the principal on whom the norm directs from and B denotes the principal on whom the norm directs to.

- $E_{A,B} \rightarrow S_{B,A}$ represents that A's past emotions influence A's current sanctions towards B.
- $S_{B,A} \rightarrow N'_{B,A}$ represents that the outcomes of A's past sanctions influence B's current decision on the outcomes of its norm.

Let us continue on the Prohibition scenario between Asher and Bella. During COVID-19, Asher wishes to stay healthy and away from infected individuals. When Asher finds that Bella violates the self-quarantine regulation in a store, he feels angry and reports Bella to the local authorities. The sanctions Bella receives would further influence her future decision to satisfy her norm towards Asher.

4.4 Pandemic Simulation Environment

We evaluate *Hermione* via a simulated pandemic environment where agents' behavior influences the spread of a pandemic. We describe the environment below.

4.4.1 Pandemic Environment

Since the outbreak of COVID-19, some policies have proved to be effective in controlling pandemic transmission, e.g., mask-wearing (Leech et al. 2021; Cheng et al. 2020), social-distancing and space ventilation (Sun and Zhai 2020), shelter-in-place (Dave et al. 2021), and lockdowns (Lau et al. 2020). We built our pandemic environment using Mesa (Masad and Kazil 2015), a Python-based framework for building, analyzing, and visualizing agent-based models.

The pandemic environment is a Markov Decision Processes (MDP), a discrete-time stochastic model, where the chosen action influences the next state. The environment includes four shared locations—home, grocery store, hospital, and park. Each home accommodates five agents. During interactions, agents move among these places.

We initialize the agents with a health state that corresponds to the Susceptible-Exposed-Infected-Recovered-Vaccinated (SEIRV) model (Yang and Wang 2020; Annas et al. 2020), beliefs and desires, and intention variables. The beliefs variable keeps the perceived and processed information from the environment and other agents. The relevant desires are a combination of preventing death and avoiding negative sanctions and a random urge to go out or stay home.

We define an action space that includes staying home, shopping, going to park, and taking the vaccine. Our reward function covers death, desire fulfilment, negative sanctions, and norm satisfaction and violation.

4.4.2 Disease Model

Viruses are constantly changing; therefore, antibodies from recovery or vaccine may not always provide complete protection. In addition, vaccines provide substantial protection against severe illness and hospitalization but much less protection against symptomatic infections. To accommodate this concern, we include the effectiveness of vaccines in the Susceptible-Exposed-Infected-Recovered-Vaccinated (SEIRV) model (Yang and Wang 2020; Annas et al. 2020). Our adopted SEIRV model separates the deceased compartment from the recovered (R), so the state transitions are complete and intuitive. As shown in Figure 4.4, each individual starts with being susceptible (S) to the virus, can be exposed (E) to the virus when that individual has any contact with an infected individual. Susceptible and exposed and recovered are healthy individuals. Being exposed to the virus, an individual can become Infectious (I). We further define Infectious as having three subclasses: asymptomatic, mild symptomatic, and critical symptomatic. After some immune response, an individual becomes either recovered or deceased from the virus. Those who recover from the virus gain antibodies that protect them from the same virus. The other way to gain protection is vaccination. Once people obtain antibodies from vaccines or recovery, they have fewer opportunities to be infected.

We base the probabilities of how COVID-19 evolves on a study in Italy (Poletti et al. 2020). According to the quantifying research, among the close contacts, 51.5% were infected. Morbidity and mortality weekly report (Thompson et al. 2021) shows that the effectiveness of mRNA vaccine with full immunization was 90% against COVID-19 infections regardless of symptom status. Since the emerging B.1.617.2 (Delta) variant is 50% more contagious than the original strain of SARS-CoV-2, we set the infection probability to 80%. With different vaccines, the effectiveness with full vaccination ranged from 67% to 93.7% among persons with the Delta variant (Lopez Bernal et al. 2021). We set the effectiveness of vaccination at 50% to speed up the simulation. Apart from vaccination, we set the probability of the symptoms to progress as Figures 4.4 and 4.5. The intuition is that each infected person provides an opportunity for the symptoms to progress to the next phase or recover. With complete vaccination, the probability of the disease evolving is halved. Table 4.3 details the transition probability. Note that the transition for healthy agents applies when coming in contact with those who are infected.

In the real world, we have a greater tendency to sanction those who severely violate norms. In our simulation, we set 50% probability to sanction those who appear mild symptoms and 80% possibility to sanction those who appear critically symptomatic but are not quarantining. Since humans have partial observability of others in the real world, we include uncertainty of observing other health states (Table 4.2). Someone who is sniffling has some probability of being perceived as symptomatic.

Table 4.2: Uncertainty on observing others' state.

Actual \ Belief	Healthy	Symptomatic	Critical Illness
Healthy	0.8	0.1	0.1
Asymptomatic	0.5	0.5	0.0
Symptomatic	0.3	0.6	0.1
Critical Illness	0.1	0.3	0.6

Table 4.3: State transition of disease with probability.

state_t \ state_{t+1}	Healthy	Asymptomatic	Mild	Critical	Deceased
Healthy	1-0.8α	0.8α	0	0	0
Asymptomatic	0.2β	1-0.36α-0.2β	0.36α	0	0
Symptomatic	0.1β	0	1-0.01α-0.1β	0.01α	0
Critical	0.05β	0	0	1-0.2α-0.05β	0.2α

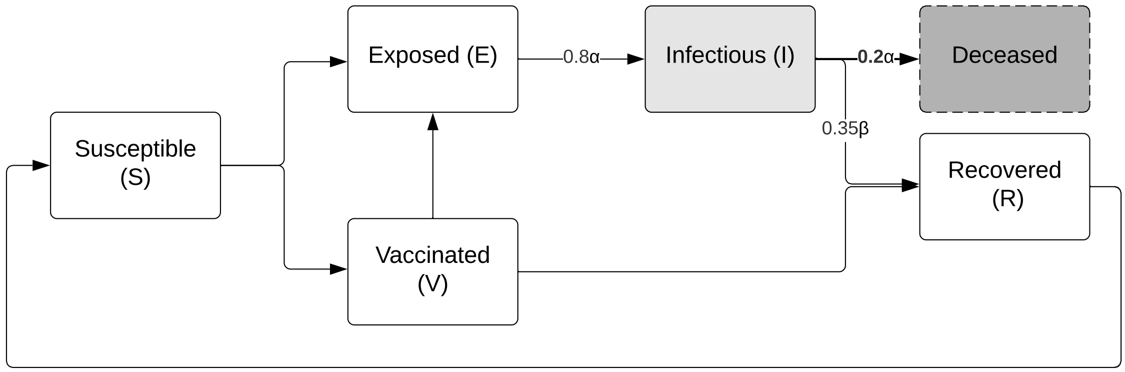


Figure 4.4: Disease model with state transition probabilities adapted from SEIRV disease model (Yang and Wang 2020; Annas et al. 2020). We set α to 0.5 for vaccinated agents and to 1.0 for unvaccinated agents; $\beta = 1.0$ for agents not staying home and $\beta = 2.0$ for agents staying home. The probability of remaining in the state is $1-$ the probability of evolving to the next state.

4.4.3 Agent Types

To answer our research questions and evaluate *Hermione*, in addition to *Hermione*, we define four types of agent societies as baselines. We describe the agents societies below:

Baseline 1: PRIMITIVE society. Agents in a primitive society do not have given social norms at the beginning. Agents obtain a negative reward when they die.

Baseline 2: SANCTIONING society. This society is based on sanctions with which agents obtain negative rewards from violating norms (Nardin et al. 2016). Agents obtain a negative reward for dying from the COVID-19.

Baseline 3: EMOTION society. This society is based on sanctions, self-directed emotions, and others' emotions (Tzeng et al. 2021). Agents consider their own and others' emotional responses to their actions in the decision-making process. In this society, agents obtain negative rewards from death or violating norms.

Baseline 4: MESSAGE society. This society is based on sanctions, that is, negative rewards for agents for violating norms. Whereas sanctions with rewards or punishments are costly in terms of cost and possible outcomes, in most cases, agents send a normative message stating what sanctions an agent will receive if it violates a norm. The normative message is adapted from (Andrighetto et al. 2013). Agents obtain negative rewards from death.

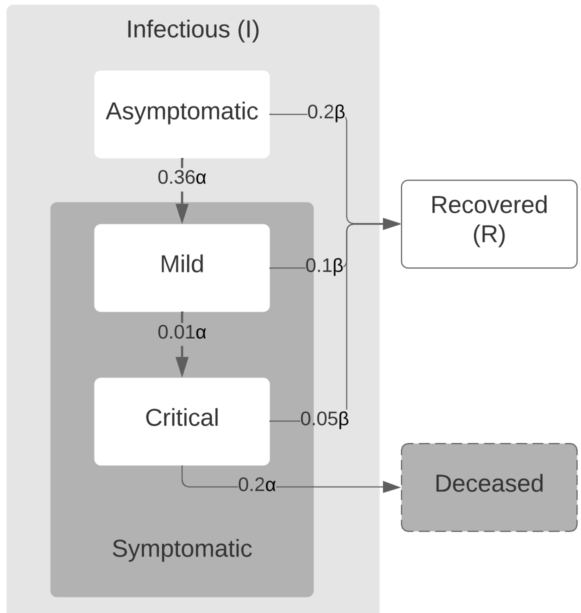


Figure 4.5: State transitions between infectious, recovered, and deceased. Infectious include three subclasses: asymptomatic, mildly symptomatic, and critically symptomatic. Numbers on the edges show the probabilities of transition. Hence, $\alpha = 0.5$ for vaccinated agents and $\alpha = 1.0$ for unvaccinated agents. And, $\beta = 1.0$ for agents not staying home and $\beta = 2.0$ for agents staying home. The probability of remaining in the state is $1-$ the probability of evolving to the next state.

***Hermione*: Emotion as information society.** Our proposed method. Agents in the *Hermione* society make decisions considering emotions and the normative information from other agents and the estimated outcome of their actions. Unlike an agent in the MESSAGE society, a *Hermione* agent reasons over emotions instead of explicit messages.

Table 4.4 summarizes the difference among our defined societies.

4.4.4 Social Expectation Modeling

Social expectations (Rummel 1975; Ajmeri et al. 2017) are a societal climate that regulates our behavior towards others and measures our perceived psychological distances from them. Specifically, social expectations are about some internalized norms in society and about how

Table 4.4: Characteristics of agent societies. Sanctioning means giving hard sanctioning at a low probability. Belief Reward indicates that incorporating believed reward from normative information as an intrinsic reward for reinforcement learning. Emotion denotes that agents consider self-directed emotion and others' other-directed emotion in decision-making. Upon receiving normative information, agents have a belief that they will be hard-sanctioned later.

Society	Sanctioning	Belief Reward	Emotion
Baseline 1. PRIMITIVE	✗	✗	✗
Baseline 2. SANCTIONING	✓	✗	✗
Baseline 3. EMOTION	✓	✗	✓
Baseline 4. MESSAGE	✓	✓	✗
Hermione	✓	✓	✓

individuals should behave. We model these expectations with social norms.

Norms. We identify the following norms.

Prohibition. Other people prohibit an infected person from staying in a public space.

$P_{quarantine}(\mathscr{A}, \mathscr{A}, \text{infected}, \text{go out of home})$

Sanction. The local authorities force an infected individual who goes out and is reported by others into home quarantine.

$S_{quarantine}(\mathscr{A}, \mathscr{A}, \text{go out of home}; \text{infected}, \text{forced stay home})$

4.5 Experiment Design

4.5.1 Hypotheses and Metrics

To address our research question $\text{RQ}_{\text{information}}$, we compute these measures:

$\mathbf{M}_{\mathbf{Healthy}}$**:** The percentage of healthy agents.

$\mathbf{M}_{\mathbf{Infected}}$**:** The percentage of agents who are infected.

$\mathbf{M}_{\mathbf{Deceased}}$**:** The percentage of deceased agents.

$\mathbf{M}_{\mathbf{Total\ infections}}$**:** The total number of infections in societies.

$\mathbf{M}_{\mathbf{Vaccinated}}$**:** The percentage of vaccinated agents.

$M_{\text{Compliance}}$: The percentage of compliance rate. Here we define the compliance of prohibition as staying at home when infected (Section 4.4.4).

$M_{\text{Forced quarantine}}$: Number of agents who are forced to stay home quarantine. Refer to Section 4.4.4, the consequent of Sanction is forced quarantine.

$M_{\text{Self-directed emotion}}$: The average self-directed emotion among agents in emotion and *Hermione* agent societies.

$M_{\text{Other-directed emotion}}$: The average other-directed emotion among agents in emotion and *Hermione* agent societies.

$M_{\text{Desire satisfaction}}$: The average desire satisfaction among agents.

To answer our research question $RQ_{\text{information}}$, we evaluate six hypotheses that correspond to the specific metric, respectively.

$H_{\text{Disease control}}$: Agent societies considering emotions have better control over disease spread compared to the societies that do not consider emotions. Specifically, infections in EMOTION and *Hermione* societies are lower than in PRIMITIVE, SANCTIONING, and MESSAGE societies. *Hermione* Agent society has more agents vaccinated than in PRIMITIVE, SANCTIONING, MESSAGE, and EMOTION societies. We compare M_{Healthy}, M_{Infected}, M_{Deceased}, $M_{\text{Total infections}}$, and $M_{\text{Vaccination}}$ to evaluate $M_{\text{Disease control}}$.

$H_{\text{Norm compliance}}$: The proportion of compliance in agent societies considering emotions is higher than the societies that do not consider emotions. Specifically, agents in EMOTION and *Hermione* societies tend to comply more than in PRIMITIVE, SANCTIONING, and MESSAGE societies. In addition, *Hermione* agent society requires less material cost to achieve stable cooperation than PRIMITIVE, SANCTIONING, EMOTION, and MESSAGE societies. We compare $M_{\text{Compliance}}$ and $M_{\text{Forced quarantine}}$ to evaluate $H_{\text{Norm compliance}}$.

H_{Emotion}: *Hermione* agent society maintains better emotion than EMOTION society. We compare $M_{\text{Self-directed emotion}}$ and $M_{\text{other-directed emotion}}$ to evaluate H_{Emotion}.

H_{Desire}: Agents in *Hermione* agent society have more desire satisfaction than PRIMITIVE, SANCTIONING, EMOTION, and MESSAGE societies. We compare $M_{\text{Desire satisfaction}}$ to evaluate H_{Desire}.

To test the statistical significance of our hypotheses, we conduct the independent t-test and measure effect size with Glass' Δ for unrelated societies (Glass 1976; Grissom and Kim 2012). We adopt Cohen's (Cohen 1988) descriptors to interpret effect size where 0.2 indicate small, 0.5 indicate medium, and 0.8 indicate large effect. An effect size less than 0.2 indicates that the difference is negligible.

4.5.2 Simulation Experiment Setup

Here we describe our simulation settings. Tables 4.5 and 4.6 show the reward function, including extrinsic rewards from the environment and intrinsic rewards from agents' internal state, in our work. With reward shaping (Ng et al. 1999), agents are provided with additional and arbitrary shaping rewards to moving towards the goal or to encourage taking some actions in some set of states. To accommodate the perceived information, we include the belief reward shaping (Marom and Rosman 2018) as another intrinsic reward in our simulation. Specifically, we incorporate beliefs of being punished in the future into our reinforcement learning simulation. Table 4.7 demonstrates how *Hermione* appraises a state.

Table 4.5: Reward function. Hard sanctioning means forced quarantine.

Component	Reward
Deceased	–2
Intention satisfaction	+1
Intention violation	–1
Hard sanctioning	–1
Norm compliance	+1
Self-directed emotion	–1
Other-directed emotion	–1

We run each simulation with 100 agents and 50 iterations for 2,000 steps. Table 4.8 lists the learning parameters. To reduce the simulation time, we choose a relatively small number of agents. To minimize deviation from coincidence, we run each simulation with 20 iterations and compute the mean values.

Table 4.6: Payoff of intentions for various actions.

Intention \ Actual	stay home	hiking	shopping	vaccination
stay home	1	–1	–1	–1
go to park	–1	1	–1	–1
go shopping	–1	–1	1	–1

Table 4.7: Emotion valence. The final valence is the combination of all that applied. Note that the other-directed emotion evaluates the action and the perceived health state of others instead of the actual health state.

Intention \ Actual	self-directed emotion	other-directed emotion
Intention satisfaction	+1	–
Intention violation	–1	–
Norm violation (self)	–1	–
Forced quarantine	–1	–
Norm satisfaction (others)	–	+1
Norm violation (others)	–	–1

Table 4.8: Learning parameters.

Parameter	Value	Comment
Learning rate	1×10^{-3}	
Discount factor	0.99	
Simulation step per action	1	
Infection %	0.3	The default fraction of infected individuals in a society

4.6 Experimental Results

We now discuss the results for our research question $\mathrm{RQ}_{\mathrm{information}}$.

Table 4.9 summarizes the simulation results and the corresponding statistical analysis for $\mathrm{RQ}_{\mathrm{information}}$.

Table 4.9: Comparing *Hermione* agent society with baseline agent societies on various metrics and their statistical analysis with Glass' Δ and p-value.

		PRIMITIVE	SANCTIONING	EMOTION	MESSAGE	*Hermione*
$H_{\text{Disease control}}$	M_{Infected}	12.6232	2.5521	0.1904	1.6778	0.1266
	p-value	< 0.001	< 0.001	0.2413	< 0.001	–
	Δ	10.1213	1.9645	0.0517	1.2564	–
	M_{Healthy}	46.3426	78.4385	98.5669	86.9903	99.1276
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	–
	Δ	44.3871	17.3975	0.4715	10.2063	–
	M_{Deceased}	41.0343	19.0094	1.2426	11.3318	0.7458
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	–
	Δ	861.9236	390.7275	10.6290	226.4754	–
	$M_{\text{Total infections}}$	48.3351	13.0024	1.1264	8.8667	0.7810
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	–
	Δ	2138.1028	549.4885	15.5287	363.5437	–
	$M_{\text{Vaccinated}}$	83.2589	27.0262	12.6629	29.3054	7.6654
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	–
	Δ	199.8857	51.1942	13.2144	57.2209	–
$H_{\text{Compliance}}$	$M_{\text{Compliance}}$	0.6162	0.9608	0.9979	0.9720	0.9971
	p-value	< 0.001	< 0.001	0.3441	< 0.001	–
	Δ	13.0792	1.2469	0.0283	0.8631	–
	$M_{\text{Forced quarantine}}$	–	0.0128	0.0004	0.0076	5×10^{-5}
	p-value	–	< 0.001	0.0156	< 0.001	–
	Δ	–	8.0678	0.2057	4.7616	–
H_{Emotion}	$M_{\text{Self-directed emotion}}$	–	–	-0.3336	–	-0.3304
	p-value	–	–	< 0.001	–	–
	Δ	–	–	0.1430	–	–
	$M_{\text{other-directed emotion}}$	–	–	0.0011	–	0.0008
	p-value	–	–	0.2339	–	–
	Δ	–	–	0.0512	–	–
H_{Desire}	M_{Desire}	0.1875	0.2667	0.3263	0.2923	0.3306
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	–
	Δ	13.6489	6.0976	0.4101	3.6534	–

4.6.1 $\mathrm{H}_{\text{Disease control}}$

To evaluate $\mathrm{H}_{\text{Disease control}}$, we measure the proportion of healthy ($\mathrm{M}_{\text{Healthy}}$), infectious ($\mathrm{M}_{\text{Infected}}$), and deceased ($\mathrm{M}_{\text{Deceased}}$) agents. In addition, we track the total number of infections ($\mathrm{M}_{\text{Total infections}}$) and the vaccination rate ($\mathrm{M}_{\text{Vaccinated}}$) in each society. Infectious agents include those who are asymptomatic, mild symptomatic, and critical symptomatic.

Figure 4.6 shows the comparison of $\mathrm{M}_{\text{Healthy}}$, $\mathrm{M}_{\text{Infected}}$, $\mathrm{M}_{\text{Deceased}}$, $\mathrm{M}_{\text{Vaccinated}}$, and $\mathrm{M}_{\text{Total infections}}$ in agent societies. We observe that, first, *Hermione* agent society has the least infected individuals (0.1266) compared to PRIMITIVE (12.6232), SANCTIONING (2.5521), EMOTION (0.1904) , and MESSAGE (1.6778) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$).

Second, *Hermione* agent society has the most healthy individuals (99.1276) compared to PRIMITIVE (46.3426), SANCTIONING (78.4385), EMOTION (98.5669) , and MESSAGE (86.9903) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$).

Third, *Hermione* agent society has the least deceased individuals (0.7458) compared to PRIMITIVE (41.0343), SANCTIONING (19.0094), EMOTION (1.2426) , and MESSAGE (11.3318) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$).

In terms of $\mathrm{M}_{\text{Total infections}}$, *Hermione* agent society has the least total number of infections (0.7810) compared to PRIMITIVE (48.3351), SANCTIONING (13.0024), EMOTION (1.1264) , and MESSAGE (8.8667) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$).

With regard to $\mathrm{M}_{\text{Vaccinated}}$, *Hermione* agent society has the least vaccination rate (7.6654) compared to PRIMITIVE (83.2589), SANCTIONING (27.0262), EMOTION (12.6629) , and MESSAGE (29.3054) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$).

4.6.2 $\mathrm{H}_{\text{Norm compliance}}$

To evaluate $\mathrm{H}_{\text{Norm compliance}}$, we measure the proportion of compliance ($\mathrm{M}_{\text{Compliance}}$) and the number of agents in forced quarantine ($\mathrm{M}_{\text{Forced quarantine}}$) in agent societies. While compliance means voluntarily staying home when infectious, forced quarantine means forced staying home by external forces. In other words, forced quarantine indicates forced compliance. Figure 4.7 exhibits plots comparing $\mathrm{M}_{\text{Compliance}}$ and $\mathrm{M}_{\text{Forced quarantine}}$. We observe that the *Hermione* agent society has a higher tendency to follow norms (0.9971) compared to PRIMITIVE (0.6162), SANCTIONING (0.9608), and MESSAGE (0.9720) societies. The differences in these results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$). Although EMOTION societies have a higher tendency of compliance (0.9979) than *Hermione* agent society (0.9971), with p-value larger than 0.05 and Glass' $\Delta \approx 0.2$, the differences in the results are not significant.

In terms of $M_{\text{Forced quarantine}}$, we notice that *Hermione* agent society has the least number of forced quarantine executed ($5e-05$) compared to SANCTIONING (0.0128), EMOTION (0.0004) , and MESSAGE (0.0076) societies. Whereas the results of statistical analysis shows that the differences between the *Hermione*, SANCTIONING, and MESSAGE societies are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$), the differences between *Hermione* and EMOTION societies are significant but small ($p < 0.05$; Glass' $\Delta \approx 0.2$).

4.6.3 H_{Emotion}

To evaluate H_{Emotion}, we measure the self-directed emotion ($M_{\text{Self-directed emotion}}$) and other-directed emotion ($M_{\text{Other-directed emotion}}$) among each society. Figure 4.8 demonstrates plots comparing $M_{\text{Self-directed emotion}}$ and $M_{\text{Other-directed emotion}}$. For self-directed emotion, we observe that *Hermione* agent society yields better emotions (-0.3304) than EMOTION society (-0.3336). However, although the differences between the *Hermione* and EMOTION societies are statistically significant ($p < 0.001$), the differences are negligible(Glass' $\Delta < 0.2$) In terms of other-directed emotion, we notice that *Hermione* agent society yields worse emotions (0.0008) than EMOTION society (0.0011). With $p > 0.05$ and Glass' $\Delta < 0.2$, the differences between the *Hermione* and the EMOTION societies are not statistically significant and negligible.

4.6.4 H_{Desire}

To evaluate H_{Desire}, we measure the desire satisfaction (M_{Desire}) in agent societies. Figure 4.9 shows the plot comparing M_{Desire} in agent societies. We observe that agents in the *Hermione* agent society have the highest desire satisfaction (0.3306) compared to the PRIMITIVE (0.1875), SANCTIONING (0.2667), EMOTION (0.3263), and MESSAGE (0.2923) societies. The differences in the results are statistically significant ($p < 0.001$; Glass' $\Delta > 0.8$ between the *Hermione*, PRIMITIVE, SANCTIONING, and MESSAGE societies. With Glass' $\Delta > 0.2$ and $p < 0.001$, the differences between the *Hermione* and the EMOTION society are statistically significant but have a small effect size.

4.6.5 Threats to Validity

First, our simulation environment has a limited action space. In reality, humans have varied reactions to regulations or situations instead of limited actions. e.g., going out with a mask, sanitizing, and cleaning themselves after going back home. All these flexible actions influence disease control. However, our focus is not to model the reality but show how normative information in form of messages and emotions can lead to better cooperation.

Second, in reality, people have different desires and preferences, thus simulating with the same desires or equal probability for desires introducing a threat of desire difference. Some people may prefer staying at home over outdoor activities. Future works could consider preference and heterogeneity among agents.

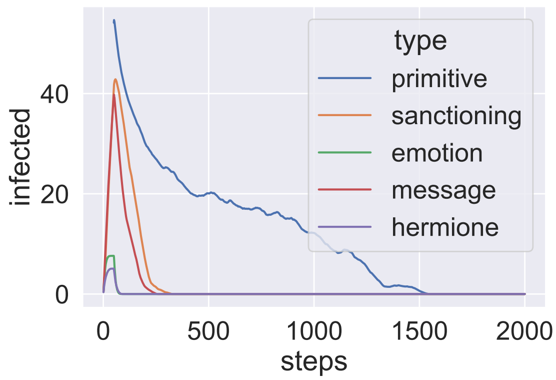
Third, to achieve data efficiency, agents in our simulation share the Q function. However, humans who have diverse preferences or values may make different decisions under the same condition. Future works could consider heterogeneity in decision making.

4.7 Conclusions and Future Works

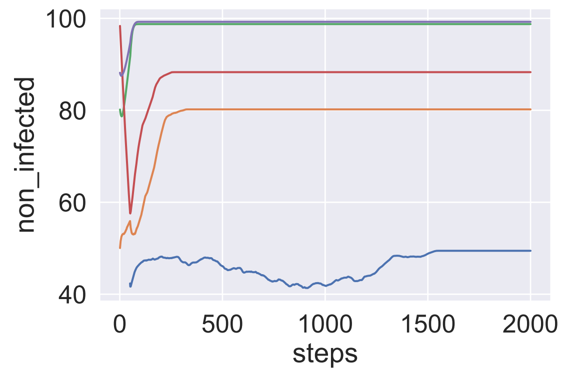
We present an agent architecture that integrates reinforcement learning with normative reasoning, cognitive architecture, and emotion modeling to answer our first research question $\mathrm{RQ}_{\mathrm{RL}}$. With reinforcement learning, agents learn from interactions and minimize human interventions and efforts. Reinforcement learning also provides an ideal space for bottom-up norms to emerge. To investigating our research question related to normative message ($\mathrm{RQ}_{\mathrm{message}}$), we simulate the COVID-19 scenario in which agents do hard sanctioning or give others warnings for norm violation. Our simulations consider three characteristics of agent societies: having hard sanctioning, incorporating prior beliefs, and emotions involved. The experiments show that with normative message given, agents cooperate better than those societies without normative details.

Apart from being part of social norms or being a sanctioning approach, emotions also transmit information. Different from the explicit message, emotions suggest a subtle attitude. We show that expressed emotion as information encourages cooperation and enforce norms.

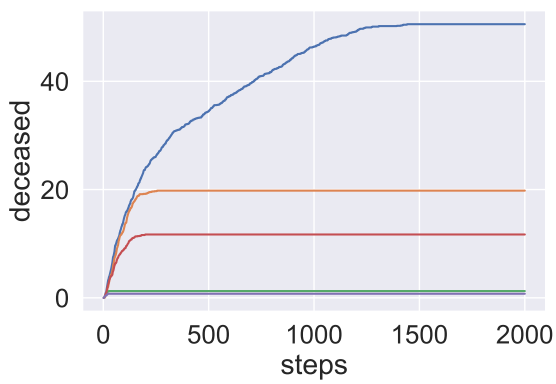
While AI has been part of our daily lives nowadays, incorporating human ethics into AI becomes a necessary problem (Murukannaiah et al. 2020; Ajmeri et al. 2020). Since human behavior is driven by the pursuit of values, studying human values helps us understand human decisions and create AI that reason along with human values (Liscio et al. 2021). Our future work includes differentiate emotions and a mix of personality in *Hermione* to provide more information and dynamic. We can investigate how different values influence human interactions in future research to achieve an individual level of heterogeneity.



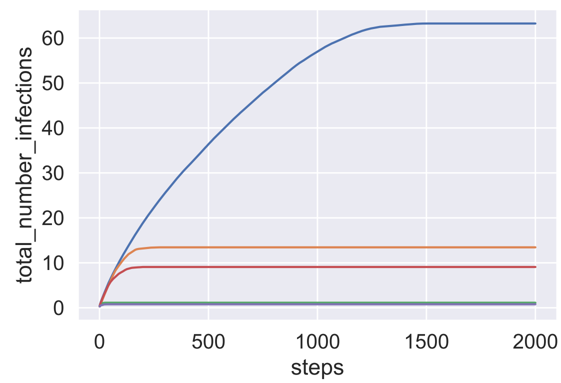
(a) The proportion of infected individuals.



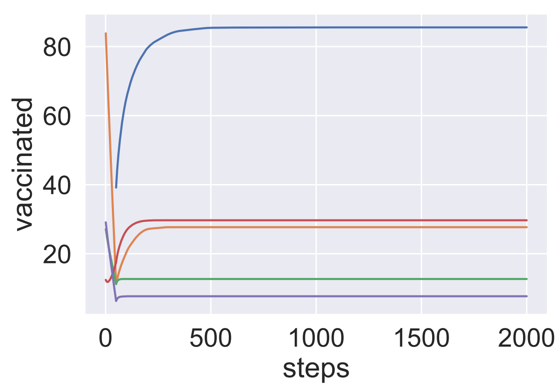
(b) The proportion of healthy individuals.



(c) The proportion of deceased individuals.

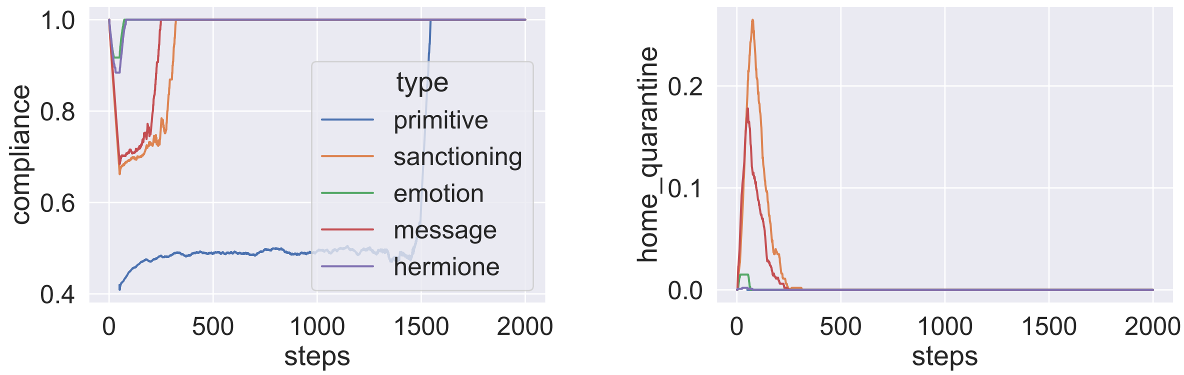


(d) Total number of infections.



(e) The proportion of vaccinated individuals.

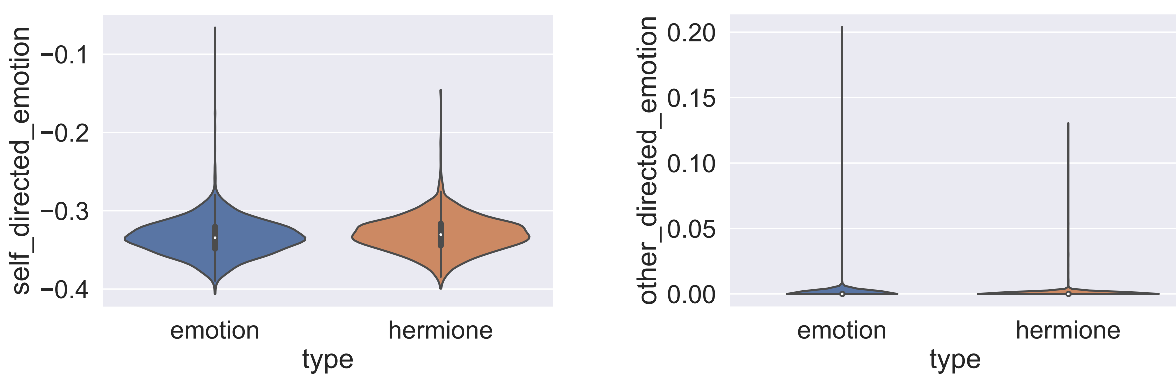
Figure 4.6: Comparing proportion of healthy(M_{Healthy}), infectious(M_{Infected}), deceased(M_{Deceased}), vaccinated ($M_{\text{Vaccinated}}$), and the total number of infections ($M_{\text{Total infections}}$) in various agent societies. The Emotion and the *Hermione* agent societies have significantly more healthy agents, less infected agents, less deceased agents, less total number of infections, and less vaccinations ($p < 0.001$; Glass' $\Delta > 0.8$) compared to the Primitive, Sanctioning, and Message societies.



(a) The proportion of compliance.

(b) The number of agents in forced quarantine.

Figure 4.7: Comparing the compliance ($M_{\text{Compliance}}$) and the number of agents in forced quarantine ($M_{\text{Forced quarantine}}$) in the *Hermione* and baseline agent societies. Compliance means voluntarily staying home when infectious. The percentage of compliance in the EMOTION and the *Hermione* agent societies is significantly higher ($p < 0.001$; Glass' $\Delta > 0.8$) than the societies that do not consider emotions. The *Hermione* agent society has significantly less agents in forced quarantine ($p < 0.001$; Glass' $\Delta > 0.8$) to achieve stable cooperation than the SANCTIONING and the MESSAGE societies. The difference between the *Hermione* and EMOTION societies is small ($p < 0.05$; Glass' $\Delta \approx 0.2$).



(a) The average self-directed emotion.

(b) The average other-directed emotion.

Figure 4.8: Comparing the average self-directed ($M_{\text{Self-directed emotion}}$) and other-directed emotion ($M_{\text{Other-directed emotion}}$) in the *Hermione* and baseline agent societies. There is no significant difference between the *Hermione* and the EMOTION societies (Glass' $\Delta < 0.2$; $p < 0.001$ for $M_{\text{Self-directed emotion}}$ and $p > 0.05$ for $M_{\text{Other-directed emotion}}$).

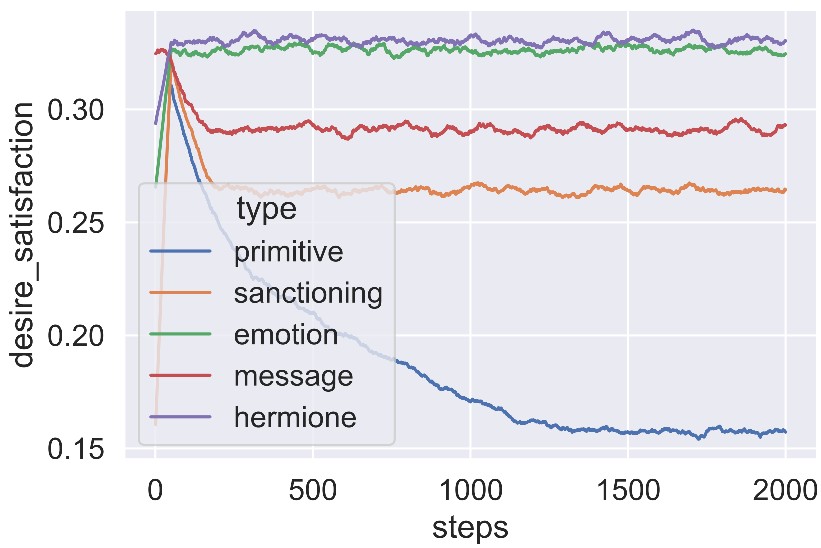


Figure 4.9: Comparing the average desire satisfaction ($M_{\text{Desire satisfaction}}$ in the *Hermione* and baseline agent societies. The *Hermione* agent society yields more desire satisfaction than baseline agent societies. In detail, the *Hermione* agent society yields significantly more desire satisfaction ($p < 0.001$; Glass' $\Delta > 0.8$) than the PRIMITIVE, SANCTIONING, and MESSAGE societies. The differences in the results of the *Hermione* and the EMOTION societies are small ($p < 0.001$; Glass' $\Delta < 0.2$).

CHAPTER

5

SOCIAL VALUES ORIENTATION FOR NORM EMERGENCE

5.1 Introduction

Cognitive is "of, relating to, being, or involving conscious intellectual activity (such as thinking, reasoning, or remembering," according to Merriam-Webster dictionary (Merriam-Webster 2021). What makes people make different decisions? Schwartz (Schwartz 2012) defined ten fundamental human values, and each of them reflects specific motivations. It's the subjective weights provided by human values, differentiate real-world decision-making.

Humans evaluate social norms based on human values. Social norms or social expectations (Rummel 1975; Ajmeri et al. 2017) are societal principles that regulate our behavior towards others via measure our perceived psychological distances from others. Most previous works related to norms did not consider human values and assumed regimented environments. However, humans do not behave in such ways. In the real world, humans with varied weights of values evaluate the outcomes of their actions subjectively and act to maximize their utility.

Example 4 ***Values.*** *Felix, who has always been a good citizen, finds someone breaks into* *his house and causes a safety threat. Being aware of the crime of assault, Felix chooses to overpower the suspect.*

However, living in a society, interacting with someone else is inevitable. Like fundamental human values, social value orientation indicates a person's preference for resource allocation between self and others. Here is an example of the real world case of SVO.

Example 5 ***SVO.*** *Elliot, an advocate of individualism and only cares about his pleasure, goes to the nightclub with Debra and Cecilia while suspicious cases exist. Although Elliot knows Debra values safety most, he chooses not to wear a mask to maximize his pleasure. As a prosocial person, Cecilia evaluates the values of everyone and chooses to wear the mask to achieve greater satisfaction between herself and others.*

With advances in technology, Artificial Intelligence (AI) has become part of our daily lives. With these modern technology, software interacts with its environment but also with each other and with humans. With humans-in-the-loop, there are emerging needs for human factors to be considered when building modern AI systems. Specifically, these AI systems should be able to reason over humans' behaviors determined by internal attitudes and external factors. AI that considers human values and sociotechnical aspects in decision-making would be more realistic and trustworthy.

While social norms define dominant behaviors in society, they may change over time. AI systems are expected to adapt to these dynamism and changing environment. Previous works on normative agents considered no basic human values nor social value orientations.

We investigate the following research questions.

$\text{RQ}_{\text{Values}}$. How does basic human values influence the emergence of norms?

RQ_{SVO}. How does social value orientation influence the emergence of norms?

To address $\text{RQ}_{\text{Values}}$, we consider basic human values within agents. These values provide a basis for state evaluation. To address RQ_{SVO}, we apply social preferences, social value orientation, along with basic human values. Our proposed framework *Fleur* combines world model, cognitive architecture, and social model. Both human values and social value orientation provide subjective weights on evaluating the world. Here we assume that agents' individual preferences are visible to other agents.

Organization

Section 5.2 describes the schematics of *Fleur*. Section 5.3 details the pandemic environment. Section 5.4 describes the simulation experiments we conduct and their results.

5.2 *Fleur*

We now discuss the schematics of *Fleur* agents.

Figure 5.1 shows the architecture of an *Fleur* agent. *Fleur* agents consists of four main components: cognitive model, world model, social model, and a decision module.

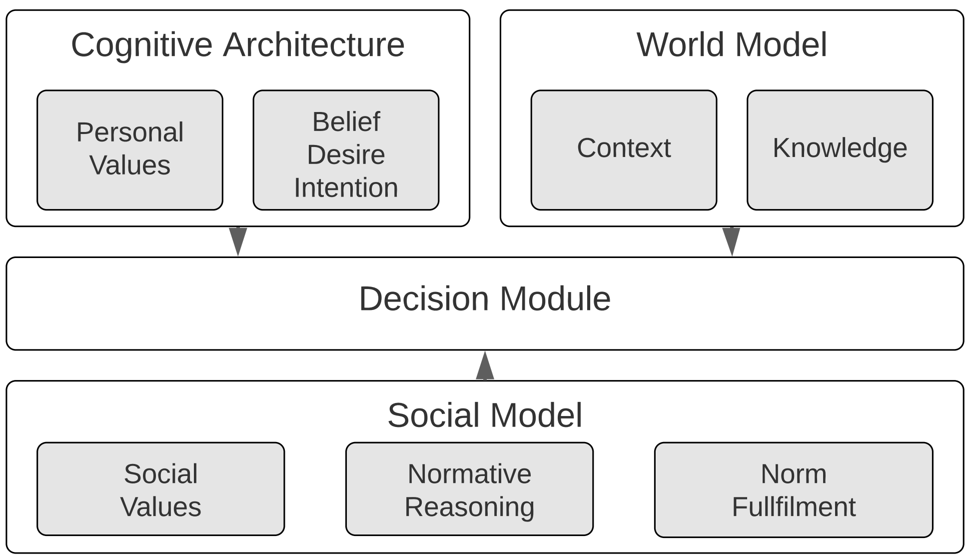


Figure 5.1: *Fleur* architecture.

5.2.1 Cognitive Model

Cognition relates to conscious intellectual activities, such as thinking, reasoning, or remembering, among which human values are essential. We introduce Schwartz's basic human values (Schwartz 2012), which defines the nature of human values and how they reflect motivations.

Figure 5.2 shows the theoretical model of human values. Human values define the intrinsic motivation of an individual and dominate how this individual thinks and evaluates everything.

We extend from the Belief-Desire-Intention (BDI) architecture (Rao and Georgeff 1991) where belief is processed information from the environment and may not reflect the actual state. Desires represent the motivations of an agent, where the intention is a plan or action to achieve the desired state an agent prefers.

Take Example 5 for instance. Elliot has an intention to have some fun at the nightclub with some COVID symptoms. However, he believes that it might be the flu and the disease of COVID is well-controlled in their area. The nightclub allows people to wear a mask of their free will. Since Elliot values pleasures more, he chooses not to wear a mask.

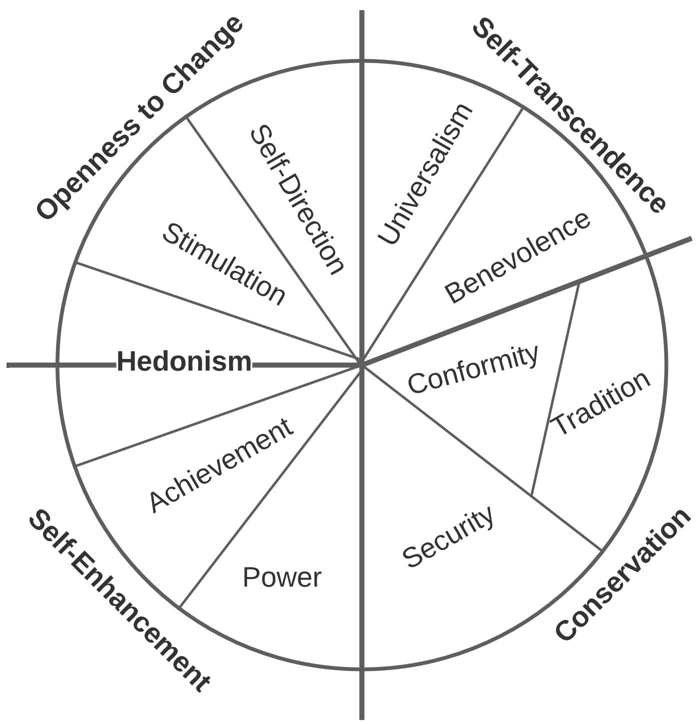


Figure 5.2: Schwartz's model of relations among motivational types of values (Schwartz 2012).

5.2.2 World Model

The world model describes the contexts in which *Fleur* agents stand and represents default knowledge set to *Fleur* agents. A context is a scenario that an agent faces. The knowledge in this model is the fact of the world. In Example 5, the context is that Cecelia is going to the nightclub and is aware of the potential probability of being infected. Under this context, Cecelia has two available actions: wear or not wear a mask. In the meantime, Cecelia is aware that a pandemic is ongoing from her knowledge.

5.2.3 Social Model

Social model includes social values, normative reasoning, and norm fulfillment. Social values define an intrinsic preference over resources allocation between agent i itself and others. Figure 5.3 demonstrates the reward distribution of different SVO types. Let $\overrightarrow{R} = (r_1, r_2, \ldots, r_n)$ represent the reward vector for a group of agents with size n. The reward for agent i is:

$$reward = w_1 \cdot r_i + w_2 \cdot r_{-i} \tag{5.1}$$

where r_i represents the reward for agent i and r_{-i} is the mean reward of all other agents interacting with agent i. w_1 and w_2 denote the weights controlling the reward distribution for self and others.

Step further for the rewards, we consider two components to the reward: (1) extrinsic reward (2) intrinsic reward. Extrinsic rewards come from the environment, while intrinsic rewards stem from human values and social preferences.

The Normative-reasoning component reasons over states, norms, and possible outcomes of satisfying or violating norms. Norm fulfillment checks if a norm has been fulfilled or violated with the selected action. Sanctions may come after Norm fulfillment. In Example 5, Cecelia is a prosocial person, which means she values her payoff (r_i) but also others' payoff (r_{-i}). In this work, we consider w_1 and w_2 as 1. Norms stem from social expectations. When most people in that nightclub wear a mask, most in the majority notice Elliot's behavior and sanction him.

5.2.4 Decision Module

The decision module generates actions based on agents' individual values and social preferences. We deploy Q-Learning (Watkins and Dayan 1992), a model-free reinforcement learning algorithm that learns from trial and error, for our agents. Q-Learning algorithm approximate the

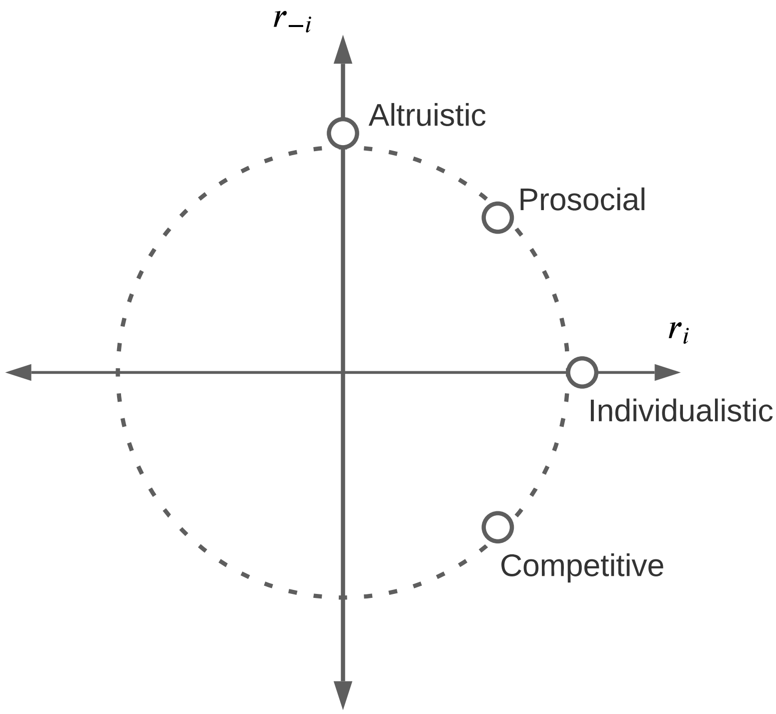


Figure 5.3: Representation of Social value orientation (Griesinger and Livingston Jr. 1973; McKee et al. 2020). r_i denotes outcome for one self while r_{-i} denotes outcomes for others.

action-state value $Q(s,a)$ (Q value), with each state and action:

$$Q'(s_t,a_t) = Q(s_t,a_t) + \alpha * (r_t + \gamma \max_{a'} Q(s_{t+1},a) - Q(s_t,a_t)) \tag{5.2}$$

where $Q'(s_t,a_t)$ represents the updated Q-value after performing action a at time t and having s_{t+1} as next state. α indicates the learning rate in the Q-value update function, and γ defines the reward discount rate. By approximating the action-state value, the Q-Learning algorithm finds the optimal policy via the expected and cumulative rewards.

Agents observe the environment, form their beliefs about the world, and update their state-value with rewards via interactions. Individual values become intrinsic rewards. During the interactions, the social model introduces feedback and considerations from other agents. SVO provides agents with different preferences of reward distribution between themselves and others. Normative reasoning enables agents to evaluate possible outcomes between complying or violating norms. Norm fulfillment brings possible sanctions.

5.3 Pandemic Environment Model

We built the pandemic environment as an intertemporal social dilemma, represented as a Markov Decision Processes (MDP). An MDP is a discrete-time stochastic model where the chosen action determines the next state and the immediate reward from state transition.

The environment includes five shared locations: home, park, grocery store, hospital, and nightclub. Agents move between these places and interact with each other. While an agent stays out of its home, it is at the risk of being infected by others if any infected individual shows up or infecting others if it is infected.

5.3.1 Disease Model

We adopt to the Susceptible-Exposed-Infected-Recovered (SEIR) model (Hethcote 2000; Yan and Liu 2006), and initiate agents with basic human values and social preferences. Refer to Figure 5.4, each individual starts with being susceptible (S) to the virus. When a susceptible individual has contact with an infected individual, that susceptible individual becomes exposed (E) to the virus. Susceptible and exposed, and recovered are considered healthy individuals. An exposed individual can become Infectious (I). We detail Infectious with three subclasses: asymptomatic, mildly symptomatic, and critical symptomatic (Figure 5.5). An infected individual can becomes either recovered or deceased from the virus. Figure 5.4 and 5.5 list the probabilities of state transition between different health states. Note that α, the parameter defining if wearing a mask or not, applies to two agents. The probability of an agent being infected is $0.5\alpha_i \times 0.5\alpha_j$ where agent i is infected and agent j is healthy.

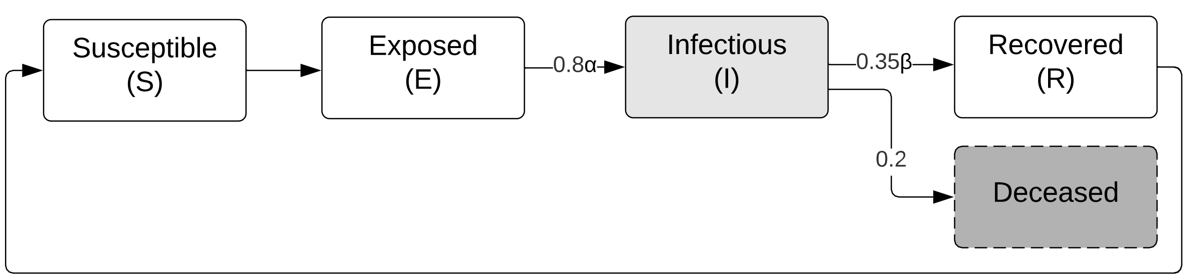


Figure 5.4: Disease model with state transition probabilities adapted from SEIR disease model (Hethcote 2000; Yan and Liu 2006). We set α to 0.5 for agents protected by wearing a mask and to 1.0 for unprotected agents; $\beta = 1.0$ for agents not staying home and $\beta = 2.0$ for agents staying home. The probability of remaining in the state is $1-$ the probability of evolving to the next state.

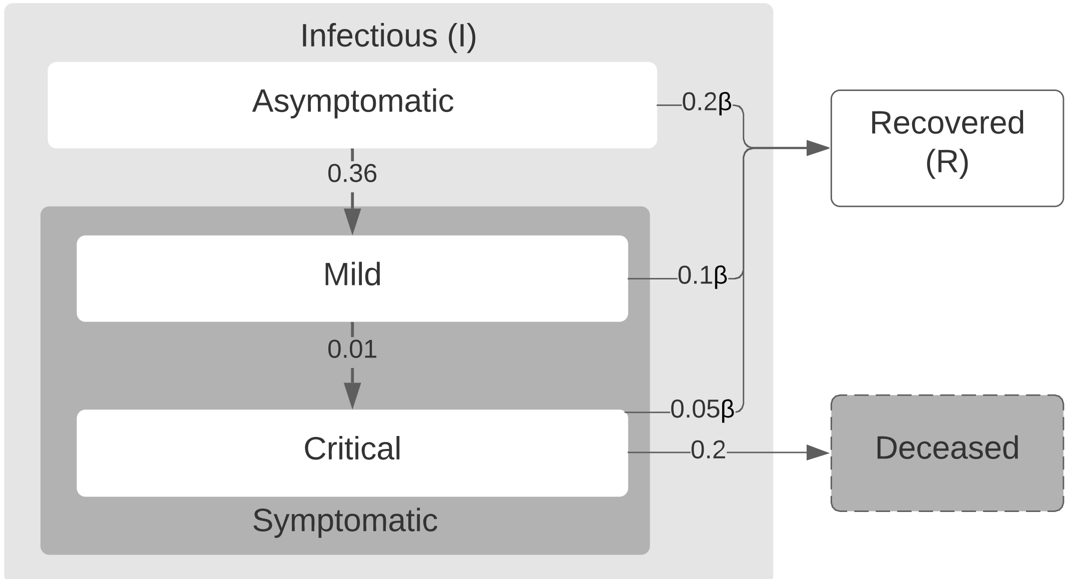


Figure 5.5: State transitions between infectious, recovered, and deceased. Infectious include three subclasses: asymptomatic, mildly symptomatic, and critically symptomatic. Numbers on the edges denote the probabilities of state transition. $\beta = 1.0$ for agents not staying home and $\beta = 2.0$ for agents staying home. The probability of remaining in the state is $1-$ the probability of evolving to the next state.

5.3.2 Social Value Orientation

Refer to Figure 5.3, suppose each action brings out self-reward S and other-reward O. An altruistic agent chooses the action that maximizes O while an individualistic agent chooses the action that maximizes S. A prosocial agent seeks to maximize both S and O while a competitive agent prefers to maximize $S - O$.

5.4 Research Plan

Considering the challenges in Example 4 and 5, we propose to address $\mathrm{RQ}_{\mathrm{Values}}$ and $\mathrm{RQ}_{\mathrm{SVO}}$ with the following simulations.

5.4.1 Experimental Setup

We propose to build our simulation with Mesa (Masad and Kazil 2015), an agent-based modeling framework in Python for creating, visualizing, and analyzing agent-based models. We will evaluate *Fleur* via a simulated pandemic environment where agents' behaviors influence the spread of a pandemic. We run the simulations on a device with 32 GB RAM and GPU NVIDIA GTX 1070 Ti.

We incorporate beliefs and desires, and intentions into our agents. An agent observes its environment and processes the perception, and then forms its beliefs about the world. Since the real world is a partially observable environment, we include uncertainty of observability. Agent observability is limited to the location where the agent stands.

In our simulation, each home accommodates one agent, and each time step hospital take care of five agents. Each episodes last 1000 steps.

Individual values. For each agent, we consider safety (security), conformity, and pleasure (hedonism) based on Schwartz's value theory (Schwartz 2012). Safety means to protect oneself. Conformity indicates compliance with social expectations. We set $2/3$ as the threshold of majority for emergence of social norms. Specifically, when more than $2/3$rd of agents in a group choose the same action, we consider such behavior as a social norm. Once social norms emerge, an agent in the majority sanctions agents in the minority. Pleasure derives from organismic needs and is the goal of hedonism. Safety in self-concern and other-concern have the same meaning. However, the pleasure for self complements the pleasure for others considering wearing or not wearing a mask. We leave out other values on account of the non-feasibility in our scenario. In terms of personal values, we generate agents with different individual value distributions and distribute these distributions equally in each group of agents.

Desires. Each step an agent has one of the five desires listed in Table 5.1 as its intention.

Beliefs. An agent's belief may not reflect the actual state. Beliefs include health state and majority behavior at a specific location. For instance, an agent may wrongly perceive an infected agent as a healthy agent. Each agent has a belief about conformity based on past observations.

Actions. Each step, an agent chooses where to go according to its intention and whether to wear a mask in a specific context.

Context. Each action at different context represent different values as in Figure 5.2.

Social Value Orientation. We consider altruistic, prosocial, individualistic, and competitive orientation in Figure 5.3.

Table 5.1: Desires table.

State	Home	Park	Grocery store	Hospital	Nightclub
Healthy	0.2	0.2	0.2	0.2	0.2
Infected	0.15	0.15	0.15	0.55	0.15

Table 5.2: Contexts and utility.

Locations	Belief	Actions	Personal Values		
			Safety	Conformity	Pleasure
Home	No potential danger	Wearing mask	0.5000	0	0
		Not wearing mask	0.5000	0	1
Park	Potential danger	Wearing mask	1	0 or 1	0.3000
		Not wearing mask	0	0 or 1	0.7000
	No potential danger	Wearing mask	0.6000	0 or 1	0.4000
		Not wearing mask	0.4000	0 or 1	0.6000
Grocery	Potential danger	Wearing mask	1.0000	0 or 1	0.5000
		Not wearing mask	0	0 or 1	0.5000
	No potential danger	Wearing mask	0.6000	0 or 1	0.5000
		Not wearing mask	0.4000	0 or 1	0.5000
Hospital	Potential danger	Wearing mask	1	0 or 1	0.5000
		Not wearing mask	0	0 or 1	0.5000
	No potential danger	Wearing mask	0.7000	0 or 1	0.5000
		Not wearing mask	0.3000	0 or 1	0.5000
Nightclub	Potential danger	Wearing mask	1.0000	0 or 1	0.3000
		Not wearing mask	0	0 or 1	0.7000
	No potential danger	Wearing mask	0.6000	0 or 1	0
		Not wearing mask	0.4000	0 or 1	1

5.4.2 Hypotheses

To address our research question $\mathrm{RQ}_{\mathrm{values}}$, we hypothesize:

$\mathbf{H}_{\mathbf{Values}}$**:** A society with more agents put pleasure in high priority has worse overall welfare.

To address our research question RQ_{SVO}, we hypothesize:

$\mathbf{H}_{\textbf{Welfare}}$: the overall A mixed preference society has better overall welfare.

$\mathbf{H}_{\textbf{Resilience}}$: A prosocial agent learns social norms faster.

5.4.3 Metrics

We compute the following metrics to address H_{Values}.

$\mathbf{M}_{\textbf{Welfare}}$ The attributes that are related to agents' welfare in a society.

- Healthy agents: The percentage of healthy agents.
- Deceased agents: The percentage of deceased agents.
- Total infections: The total number of infections.

$\mathbf{M}_{\textbf{Compliance}}$ The percentage of compliance. Here we define compliance as following social norms.

To address $\text{H}_{\text{Welfare}}$, we compute the following metrics.

$\mathbf{M}_{\textbf{Self reward}}$ The average payoff for self.

$\mathbf{M}_{\textbf{Others' reward}}$ The average payoff for others.

To address $\text{H}_{\text{Resilience}}$, we compute the following metrics.

$\mathbf{M}_{\textbf{Resilience}}$ How soon does a society return to a good state.

5.4.4 Experiment: Diverse Agent Society

To address $\text{RQ}_{\text{Values}}$ with H_{Values}, we propose to train 40 heterogeneous agents with individualistic orientation. We hypothesize that a agent society with more agents put pleasure in high priority has worse overall welfare. We then compute $\text{M}_{\text{Welfare}}$ and evaluate our hypothesis.

5.4.5 Experiment: Diverse Agent Society vs Single Orientation Agent Society

To address RQ_{SVO} with $H_{Welfare}$, we define five societies as below and train 40 heterogeneous agents for each society.

Altruistic society A society of agents with altruistic orientation.

Prosocial society A society of agents with prosocial orientation.

Selfish society A society of agents with individualistic orientation.

Competitive society A society of agents with competitive orientation.

Distributed society A society with four types of SVO equally distributed.

We then compute $M_{Self\ reward}$ and $M_{Other\ reward}$ and evaluate $H_{Welfare}$.

5.4.6 Experiment: Newcomer

To address RQ_{SVO} with $H_{Resilience}$, we propose to train a new learning agent with the trained agents in a diverse agent society. Our hypothesis is that an agent society that yields greater resilience requires less time for the newcomer to learn norms. We then compute $M_{Resilience}$ and evaluate $H_{Resilience}$.

CHAPTER

6

DELIBERATE JUSTIFICATION

6.1 Introduction

While humans evaluate social norms based on their values, they are flexible in accepting exceptions. Consider a self-driving car scenario. In the general case, all vehicles have to follow the traffic rules. However, when one vehicle is on the duty of emergency patient transport, other vehicles aware of this fact will accept it as an exception to traffic rules. Previous work on conflict resolution reveal all information to others. Yet, some information may be sensitive in the real world, or someone has a preference not to share the information. Humans have the capacity for this kind of flexibility. Therefore, we propose work on modeling flexible justifications and study how this practice influences norm emergence.

6.1.1 Hypothesis

$\mathbf{H}_{\mathbf{Justification}}$**:** Deliberate justifications increase general satisfaction without sacrificing agents' values.

6.2 Research Plan

We propose a value-based framework that will evaluate the state based on human values. We will experiment on a scenario that agents have to decide what information to share deliberately.

REFERENCES

Nirav Ajmeri, Pradeep K. Murukannaiah, Hui Guo, and Munindar P. Singh. Arnor: Modeling social intelligence via norms to engineer privacy-aware personal agents. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 230–238, São Paulo, May 2017. IFAAMAS.

Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. Robust norm emergence by revealing and reasoning about context: Socially intelligent agents for enhancing privacy. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 28–34, Stockholm, July 2018. IJCAI. doi: 10.24963/ijcai.2018/4.

Nirav Ajmeri, Hui Guo, Pradeep K. Murukannaiah, and Munindar P. Singh. Elessar: Ethics in norm-aware agents. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems, (AAMAS)*, pages 16–24, Auckland, May 2020. IFAAMAS. doi: 10.5555/3398761.3398769.

Shams Al-Amin, Nirav Ajmeri, Hongying Du, Emily Z. Berglund, and Munindar P. Singh. Toward effective adoption of secure software development practices. *Simulation Modelling Practice and Theory (SIMPAT)*, 85:33–46, June 2018. doi: 10.1016/j.simpat.2018.03.006.

Bexy Alfonso Espinosa. *Agents with Affective Traits for Decision-Making in Complex Environments*. PhD thesis, Universitat Politècnica de València, 2017.

Linda J.S. Allen. Some discrete-time SI, SIR, and SIS Epidemic Models. *Mathematical Biosciences*, 124(1):83–105, 1994.

Giulia Andrighetto, Jordi Brandts, Rosaria Conte, Jordi Sabater-Mir, Hector Solaz, and Daniel Villatoro. Punish and voice: Punishment enhances cooperation when combined with norm-signalling. *PLOS ONE*, 8(6):e64941, 2013.

Suwardi Annas, Muh Isbar Pratama, Muh Rifandi, Wahidah Sanusi, and Syafruddin Side. Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia. *Chaos, Solitons & Fractals*, 139:110072, 2020.

Estefania Argente, Elena Del Val, Daniel Perez-Garcia, and Vicente Botti. Normative emotional agents: A viewpoint paper. *IEEE Transactions on Affective Computing*, 2020.

Simon Baron-Cohen, Alan M. Leslie, and Uta Frith. Does the autistic child have a "theory of mind"? *Cognition*, 21(1):37–46, 1985.

Mathieu Bourgais, Patrick Taillandier, and Laurent Vercouter. Ben: An agent architecture for explainable and expressive behavior in social simulation. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, pages 147–163. Springer, 2019.

Michael E. Bratman. *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, Massachusetts, 1987.

Jan Broersen, Mehdi Dastani, Joris Hulstijn, Zisheng Huang, and Leendert van der Torre. The BOID architecture: Conflicts between beliefs, obligations, intentions and desires. In *Proceedings of the 5th International Conference on Autonomous Agents*, pages 9–16, 2001.

Vincent Chi-Chung Cheng, Shuk-Ching Wong, Vivien Wai-Man Chuang, Simon Yung-Chun So, Jonathan Hon-Kwan Chen, Siddharth Sridhar, Kelvin Kai-Wang To, Jasper Fuk-Woo Chan, Ivan Fan-Ngai Hung, Pak-Leung Ho, and Kwok-Yung Yuen. The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *Journal of Infection*, 81(1):107–114, 2020.

Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 2nd edition, 1988.

Dhaval Dave, Andrew I. Friedson, Kyutaro Matsuzawa, and Joseph J. Sabia. When do shelter-in-place orders fight COVID-19 best? Policy heterogeneity across states and adoption time. *Economic Inquiry*, 59(1):29–52, 2021.

Jan de Mooij, Davide Dell'Anna, Parantapa Bhattacharya, Mehdi Dastani, Brian Logan, and Samarth Swarup. Quantifying the effects of norms on COVID-19 cases using an agent-based simulation. In *Proceedings of the The 22nd International Workshop on Multi-Agent-Based Simulation (MABS)*, 2021.

Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz. Runtime revision of norms and sanctions based on agent preferences. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1609–1617. International Foundation for Autonomous Agents and Multiagent Systems, 2019.

Davide Dell'Anna, Mehdi Dastani, and Fabiano Dalpiaz. Runtime revision of sanctions in normative multi-agent systems. *Autonomous Agents and MultiAgent Systems (JAAMAS)*, 34 (2):43, 2020.

Frank Dignum, Virginia Dignum, Paul Davidsson, Amineh Ghorbani, Mijke van der Hurk, Maarten Jensen, Christian Kammler, Fabian Lorig, Luis Gustavo Ludescher, Alexander Melchior, René Mellema, Cezara Pastrav, Loïs Vanhee, and Harko Verhagen. Analysing the combined health, social and economic impacts of the corovanvirus pandemic using agent-based social simulation. *Minds and Machines*, 30(2):177–194, 2020.

Marco D'Orazio, Gabriele Bernardini, and Enrico Quagliarini. Sustainable and resilient strategies for touristic cities against COVID-19: An agent-based approach. *Safety Science*, 142: 105399, 2021.

Ward Edwards. The theory of decision making. *Psychological Bulletin*, 51(4):380, 1954.

Christopher Frantz and Gabriella Pigozzi. Modeling norm dynamics in multiagent systems. *FLAP*, 5(2):491–564, 2018.

Nico H. Frijda. *The Emotions*. Cambridge University Press, 1986.

Adam Gerace. Internal and external attributions. In *Encyclopedia of Personality and Individual Differences*, pages 2328–2334. Springer, 2020.

Gene V. Glass. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5 (10):3–8, 1976.

Donald W. Griesinger and James W. Livingston Jr. Toward a model of interpersonal motivation in experimental games. *Behavioral Science*, 18(3):173–188, 1973.

Robert J. Grissom and John J. Kim. *Effect Sizes for Research: Univariate and Multivariate Applications*. Routledge, Abingdon-on-Thames, 2012.

Jianye Hao, Jun Sun, Guangyong Chen, Zan Wang, Chao Yu, and Zhong Ming. Efficient and robust emergence of norms through heuristic collective learning. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 12(4):1–20, 2017.

Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

Christopher D. Hollander and Annie S. Wu. The current state of normative agent-based systems. *Journal of Artificial Societies and Social Simulation*, 14(2):6, 2011.

Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. Revani: Revising and verifying normative specifications for privacy. *IEEE Intelligent Systems (IS)*, 31(5):8–15, September 2016. doi: 10.1109/MIS.2016.89.

Anup K. Kalia, Nirav Ajmeri, Kevin Chan, Jin-Hee Cho, Sibel Adalı, and Munindar P. Singh. The interplay of emotions and norms in multiagent systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 371–377, Macau, August 2019. IJCAI. doi: 10.24963/ijcai.2019/53.

Dacher Keltner and Jonathan Haidt. Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13(5):505–521, 1999.

Hien Lau, Veria Khosrawipour, Piotr Kocbach, Agata Mikolajczyk, Justyna Schubert, Jacek Bania, and Tanja Khosrawipour. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of Travel Medicine*, 27(3):taaa037, 2020.

Averill M. Law, W. David Kelton, and W. David Kelton. *Simulation modeling and analysis*, volume 3. McGraw-Hill New York, 2000.

Richard S. Lazarus. Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46(8):819, 1991.

Gavin Leech, Charlie Rogers-Smith, Jonas Benjamin Sandbrink, Benedict Snodin, Robert Zinkov, Benjamin Rader, John S. Brownstein, Yarin Gal, Samir Bhatt, Mrinank Sharma, Sören Mindermann, Jan Markus Brauner, and Laurence Aitchison. Mass mask-wearing notably reduces COVID-19 transmission. *medRxiv*, 2021.

Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, and Pradeep K. Murukannaiah. Axies: Identifying and evaluating context-specific values. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 799–808, 2021.

Changliu Liu. A microscopic epidemic model and pandemic prediction using multi-agent reinforcement learning. *arXiv preprint arXiv:2004.12959*, 2020.

Jamie Lopez Bernal, Nick Andrews, Charlotte Gower, Eileen Gallagher, Ruth Simmons, Simon Thelwall, Julia Stowe, Elise Tessier, Natalie Groves, Gavin Dabrera, Richard Myers, Colin N.J. Campbell, Gayatri Amirthalingam, Matt Edmunds, Maria Zambon, Kevin E. Brown, Susan Hopkins, Meera Chand, and Mary Ramsay. Effectiveness of COVID-19 vaccines against the B. 1.617. 2 (Delta) variant. *New England Journal of Medicine*, 2021.

Ofir Marom and Benjamin Rosman. Belief reward shaping in reinforcement learning. In *Proceedings of the 32nd Conference on Artificial Intelligence (AAAI)*, pages 3762–3769, New Orleans, 2018. AAAI Press.

Stacy C. Marsella and Jonathan Gratch. Ema: A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1):70–90, 2009.

David Masad and Jacqueline Kazil. MESA: An agent-based modeling framework. In *Proceedings of the 14th PYTHON in Science Conference*, pages 53–60, 2015.

Kevin R. McKee, Ian M. Gemp, Brian McWilliams, Edgar A. Duéñez-Guzmán, Edward Hughes, and Joel Z. Leibo. Social diversity and social preferences in mixed-motive reinforcement learning. In Amal El Fallah Seghrouchni, Gita Sukthankar, Bo An, and Neil Yorke-Smith, editors, *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 869–877. International Foundation for Autonomous Agents and Multiagent Systems, 2020.

Merriam-Webster. Cognitive. In *Merriam-Webster.com dictionary*. Encyclopædia Britannica, Inc., October 2021. URL `https://www.merriam-webster.com/dictionary/cognitive`.

Stanley Milgram, Hilary J. Liberty, Raymond Toledo, and Joyce Wackenhut. Response to intrusion into waiting lines. *Journal of Personality and Social Psychology*, 51(4):683, 1986.

Thomas M. Moerland, Joost Broekens, and Catholijn M. Jonker. Emotion in reinforcement learning agents and robots: A survey. *Machine Learning*, 107(2):443–480, 2018.

Andreasa Morris-Martin, Marina De Vos, and Julian Padget. Norm emergence in multiagent systems: A viewpoint paper. *Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 33(6): 706–749, 2019.

Partha Mukherjee, Sandip Sen, and Stéphane Airiau. Norm emergence under constrained interactions in diverse societies. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 779–786, 2008.

Pradeep K. Murukannaiah, Nirav Ajmeri, Catholijn M. Jonker, and Munindar P. Singh. New foundations of ethical multiagent systems. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1706–1710, Auckland, May 2020. IFAAMAS. doi: 10.5555/3398761.3398958. Blue Sky Ideas Track.

Luis G. Nardin, Tina Balke-Visser, Nirav Ajmeri, Anup K. Kalia, Jaime S. Sichman, and Munindar P. Singh. Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *The Knowledge Engineering Review (KER)*, 31(2):142–166, March 2016. doi: 10.1017/S0269888916000023.

Andrew Y. Ng, Daishi Harada, and Stuart Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proceedings of the 16th International Conference on Machine Learning (ICML)*, volume 99, pages 278–287, 1999.

Barbara Nussbaumer-Streit, Verena Mayr, Andreea Iulia Dobrescu, Andrea Chapman, Emma Persad, Irma Klerings, Gernot Wagner, Uwe Siebert, Dominic Ledinger, Casey Zachariah, and Gerald Gartlehner. Quarantine alone or in combination with other public health measures to control COVID-19: A rapid review. *Cochrane Database of Systematic Reviews*, 4(4), 2020.

Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, New York, 1988.

Piero Poletti, Marcello Tirani, Danilo Cereda, Filippo Trentini, Giorgio Guzzetta, Giuliana Sabatino, Valentina Marziano, Ambra Castrofino, Francesca Grosso, Gabriele Del Castillo, Raffaella Piccarreta, ATS Lombardy COVID-19 Task Force, Aida Andreassi, Alessia Melegaro, Maria Gramegna, Marco Ajelli, and Stefano Merler. Probability of symptoms and critical disease after SARS-CoV-2 infection. *arXiv preprint arXiv:2006.08471*, 2020.

Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI-architecture. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, 1991.

Rudolph J. Rummel. Understanding conflict and war: Vol. 1: The dynamic psychological field. *Sage Publications*, 1975.

Bastin Tony Roy Savarimuthu and Stephen Cranefield. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multiagent and Grid Systems*, 7(1):21–54, 2011.

Shalom H. Schwartz. An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):2307–0919, 2012.

Norbert Schwarz. Emotion, cognition, and decision making. *Cognition and Emotion*, 14(4): 433–440, 2000.

Sandip Sen and Stephane Airiau. Emergence of norms through social learning. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*, volume 1507, page 1512, 2007.

Herbert A. Simon. *The New Science of Management Decision*. Harper & Brothers, 1960.

Munindar P. Singh. Norms as a basis for governing sociotechnical systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):21:1–21:23, December 2013. doi: 10.1145/2542182.2542203.

Chanjuan Sun and Zhiqiang Zhai. The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable Cities and Society*, page 102390, 2020.

Mark G. Thompson, Jefferey L. Burgess, Allison L. Naleway, Harmony L. Tyner, Sarang K. Yoon, Jennifer Meece, Lauren E.W. Olsho, Alberto J. Caban-Martinez, Ashley Fowlkes, Karen Lutrick, Jennifer L. Kuntz, Kayan Dunnigan, Marilyn J. Odean, Kurt T. Hegmann, Elisha Stefanski, Laura J. Edwards, Natasha Schaefer-Solle, Lauren Grant, Katherine Ellingson, ..., and Manjusha Gaglani. Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—eight US locations, December 2020–March 2021. *Morbidity and Mortality Weekly Report*, 70(13):495, 2021.

Sz-Ting Tzeng, Nirav Ajmeri, and Munindar P. Singh. Noe: Norms emergence and robustness based on emotions in multiagent systems. *arXiv preprint arXiv:2104.15034*, 2021.

Christian von Scheve, Daniel Moldt, Julia Fix, and Rolf von Luede. My agents love to conform: Norms and emotion in the micro-macro link. *Computational and Mathematical Organization Theory*, 12(2-3):81–100, 2006.

C. Jason Wang, Chun Y. Ng, and Robert H. Brook. Response to COVID-19 in Taiwan: Big data analytics, new technology, and proactive testing. *JAMA*, 323(14):1341–1342, 2020.

Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.

Heinz Wimmer and Josef Perner. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1):103–128, 1983.

Yang Wu and Laura E Schulz. Understanding social display rules: Using one person's emotional expressions to infer the desires of another. *Child Development*, 91(5):1786–1799, 2020.

Yang Wu, Chris L. Baker, Joshua B. Tenenbaum, and Laura E Schulz. Rational inference of beliefs and desires from emotional expressions. *Cognitive Science*, 42(3):850–884, 2018a.

Yang Wu, Jennah Haque, and Laura Schulz. Children can use others' emotional expressions to infer their knowledge and predict their behaviors in classic false belief tasks. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society (CogSci)*, 2018b.

Ping Yan and Shengqiang Liu. SEIR epidemic model with delay. *The ANZIAM Journal*, 48(1): 119–134, 2006.

Chayu Yang and Jin Wang. A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Mathematical Biosciences and Engineering MBE*, 17(3):2708, 2020.

APPENDICES